RESEARCH CENTER

FIELD

Activity Report 2014

Section New Results

Edition: 2015-03-24

ABS Project-Team

5. New Results

5.1. Highlights of the Year

In 2014, two achievements are worth noticing:
Analysis of large assemblies using native mass spectrometry data. Native mass spectrometry is about to revolutionize structural biology, since such experiments give access to the composition in terms of subunits of large macro-molecular assemblies, usually beyond reach for classical experimental techniques. In this context, we designed an algorithm to infer pairwise contacts within subunits of large macro-molecular assemblies – see section 5.3.1 . To the best of our knowledge, our algorithm is the only one whose performances can be precisely analyzed, the contenders being of heuristic nature.

Analysis and comparison of conformational ensembles and sampled energy landscapes. A key property governing the behavior of many biophysical systems is the classical enthalpy - entropy balance, which is the root of thermodynamics. Therefore, studying the way a protein folds or the way two proteins assemble requires unveiling properties of ensembles of conformations of the system scrutinized. In this context, we designed novel methods to analyze and compare collections of conformations and the associated energy landscape – see section 5.4.1 . The algorithms are based on state-of-the-art techniques from computational topology (Morse theory, Morse homology), and optimal transportation.

5.2. Modeling Interfaces and Contacts

Docking, scoring, interfaces, protein complexes, Voronoi diagrams, arrangements of balls.

The work undertaken in this vein in 2014 will be finalized in 2015.

5.3. Modeling Macro-molecular Assemblies

Macro-molecular assembly, reconstruction by data integration, proteomics, modeling with uncertainties, curved Voronoi diagrams, topological persistence.

5.3.1. Connectivity Inference in Mass Spectrometry based Structure Determination

Participants: Frédéric Cazals, Deepesh Agarwal.

In collaboration with C. Caillouet, and D. Coudert, from the COATI project-team (Inria - I3S (CNRS, University of Nice Sophia Antipolis)).

Consider a set of oligomers listing the subunits involved in sub-complexes of a macro-molecular assembly, obtained e.g. using native mass spectrometry or affinity purification. Given these oligomers, connectivity inference (CI) consists of finding the most plausible contacts between these subunits, and minimum connectivity inference (MCI) is the variant consisting of finding a set of contacts of smallest cardinality. MCI problems avoid speculating on the total number of contacts, but yield a subset of all contacts and do not allow exploiting a priori information on the likelihood of individual contacts.

In this paper [14], we present two novel algorithms, MILP-W and MILP-WB. The former solves the minimum weight connectivity inference (MWCI), an optimization problem whose criterion mixes the number of contacts and their likelihood. The latter uses the former in a bootstrap fashion, to improve the sensitivity and the specificity of solution sets.

Experiments on three systems (yeast exosome, yeast proteasome lid, human eiF3), for which reference contacts are known (crystal structure, cryo electron microscopy, cross-linking), show that our algorithms predict contacts with high specificity and sensitivity, yielding a very significant improvement over previous work, typically a twofold increase in sensitivity.

The software accompanying this paper is made available, and should prove of ubiquitous interest whenever connectivity inference from oligomers is faced.

5.4. Modeling the Flexibility of Macro-molecules

Protein, flexibility, collective coordinate, conformational sampling dimensionality reduction.

5.4.1. Conformational Ensembles and Sampled Energy Landscapes: Analysis and Comparison

Participants: Frédéric Cazals, Tom Dreyfus, Christine Roth.

In collaboration with D. Mazauric (Inria Sophia Antipolis Méditerranée, Geometrica) and C. Robert (IBPC / CNRS, Paris).

In this work, we present novel algorithms and software addressing four core problems in computational structural biology, namely analyzing a conformational ensemble, comparing two conformational ensembles, analyzing a sampled energy landscape, and comparing two sampled energy landscapes [15]. Using recent developments in computational topology, graph theory, and combinatorial optimization, we make two notable contributions. First, we present a generic algorithm analyzing height fields. We then use this algorithm to perform density based clustering of conformations, and to analyze a sampled energy landscape in terms of basins and transitions between them. In both cases, topological persistence is used to manage (geometric) frustration. Second, we introduce two algorithms to compare transition graphs. The first is the classical *earth mover distance* metric which depends only on local minimum energy configurations along with their statistical weights, while the second incorporates topological constraints inherent to conformational transitions.

Illustrations are provided on a simplified protein model (BLN69), whose frustrated potential energy landscape has been thoroughly studied.

The software implementing our tools is also made available, and should prove valuable wherever conformational ensembles and energy landscapes are used.

5.5. Algorithmic Foundations

Computational geometry, Computational topology, Voronoi diagrams, α-shapes, Morse theory.

5.5.1. Mass Transportation Problems with Connectivity Constraints

Participant: Frédéric Cazals.

In collaboration with D. Mazauric (Inria Sophia Antipolis Méditerranée, Geometrica).

Given two graphs, the supply and the demand graphs, we analyze the mass transportation problem between their vertices, under connectivity constraints [16]. More precisely, for every subset of supply nodes inducing a connected component of the supply graph, we require that the set of demand nodes receiving non-zero flow from this subset induces a connected component of the demand graph. As opposed to the classical problem, a.k.a the earth mover distance (EMD), which is amenable to linear programming (LP), this new problem is very difficult to solve, and we make four contributions. First, we formally introduce two optimal transportation problems, namely minimum-cost flow under connectivity constraints problem (EMD-CC) and maximum-flow under cost and connectivity constraints problem (EMD-CCC). Second, we prove that the decision version of EMD-CC is NP-complete even for very simple classes of instances. We deduce that the decision version of EMD-CCC is NP-complete, and also prove that EMD-CC is not in APX even for simple classes of instances. Third, we develop a greedy heuristic algorithm returning admissible solutions, of time complexity $O(n^3m^2)$ with n and m the numbers of vertices of the supply and demand graphs, respectively. Finally, on the experimental side, we compare the transport plans computed by our greedy method against those produced by the aforementioned LP. Using synthetic landscapes (Voronoi landscapes), we show that our greedy algorithm is effective for graphs involving up to 1000 nodes. We also show the relevance of our algorithms to compare energy landscapes of biophysical systems (protein models).

5.5.2. *Ciruvis: a web-based Tool for Rule Networks and Interaction Detection using Rule-based Classifiers*

Participant: Simon Marillet.

In collaboration with J. Komorowski and S. Bornelöv (Uppsala University).

The use of classification algorithms is becoming increasingly important for the field of computational biology. However, not only the quality of the classification, but also its biological interpretation is important. This interpretation may be eased if interacting elements can be identified and visualized, something that requires appropriate tools and methods.

We developed a new approach to detecting interactions in complex systems based on classification [12]. Using rule-based classifiers, we previously proposed a rule network visualization strategy that may be applied as a heuristic for finding interactions. We now complement this work with Ciruvis, a web-based tool for the construction of rule networks from classifiers made of IF-THEN rules. Simulated and biological data served as an illustration of how the tool may be used to visualize and interpret classifiers. Furthermore, we used the rule networks to identify feature interactions, compared them to alternative methods, and computationally validated the findings. Rule networks enable a fast method for model visualization and provide an exploratory heuristic to interaction detection. The tool is made freely available on the web and may thus be used to aid and improve rule-based classification.

ALF Project-Team

6. New Results

6.1. Highlights of the Year

André Seznec and Pierre Michaud won the 4th Championship Branch Prediction in all the 3 categories, 4KB, 32 KB and unlimited storage predictors [23], [33], thus confirming the past championships in 2011, 2006 and 2004.

6.2. Processor Architecture

Participants: Pierre Michaud, Bharath Narasimha Swamy, Sylvain Collange, Erven Rohou, André Seznec, Arthur Perais, Surya Khizakanchery Natarajan, Sajith Kalathingal, Tao Sun, Andrea Mondelli, Aswinkumar Sridharan, Alain Ketterlin.

Processor, cache, locality, memory hierarchy, branch prediction, multicore, power, temperature

Multicore processors have now become mainstream for both general-purpose and embedded computing. Instead of working on improving the architecture of the next generation multicore, with the DAL project, we deliberately anticipate the next few generations of multicores. While multicores featuring 1000s of cores might become feasible around 2020, there are strong indications that sequential programming style will continue to be dominant. Even future mainstream parallel applications will exhibit large sequential sections. Amdahl's law indicates that high performance on these sequential sections is needed to enable overall high performance on the whole application. On many (most) applications, the effective performance of future computer systems using a 1000-core processor chip will significantly depend on their performance on both sequential code sections and single threads.

We envision that, around 2020, the processor chips will feature a few complex cores and many (maybe 1000's) simpler, more silicon and power effective cores.

In the DAL research project, http://www.irisa.fr/alf/dal, we explore the microarchitecture techniques that will be needed to enable high performance on such heterogeneous processor chips. Very high performance will be required on both sequential sections, -legacy sequential codes, sequential sections of parallel applications-, and critical threads on parallel applications, -e.g. the main thread controlling the application. Our research focuses essentially on enhancing single process performance.

6.2.1. Microarchitecture

6.2.1.1. Branch prediction

Participants: André Seznec, Pierre Michaud.

We submitted 3 predictors to the 4th Championship Branch Prediction that took place along with the ISCA 2014 conference [33], [22], [23]. Our predictors combine some branch prediction techniques that we introduced in our previous works, in particular TAGE [10] and GEHL [12]. The predictor we submitted to the 4KB and 32KB tracks was ranked first [33] in both tracks. The 3 predictors we submitted to the unlimited-size track took the first three ranks. We have established a new reference point for branch predictability limits [23].

The 12 competing predictors were mostly using already published branch prediction techniques. The main learning from this year's contest is that choosing the right combination of techniques for the given constraints is at least as important as trying to specialize branch predictors for certain branch behaviors.

6.2.1.2. Revisiting Value Prediction

Participants: Arthur Perais, André Seznec.

Value prediction was proposed in the mid 90's to enhance the performance of high-end microprocessors. The research on Value Prediction techniques almost vanished in the early 2000's as it was more effective to increase the number of cores than to dedicate some silicon area to Value Prediction. However high end processor chips currently feature 8-16 high-end cores and the technology will allow to implement 50-100 of such cores on a single die in a foreseeable future. Amdahl's law suggests that the performance of most workloads will not scale to that level. Therefore, dedicating more silicon area to value prediction in high-end cores might be considered as worthwhile for future multicores.

First, we introduce a new value predictor VTAGE harnessing the global branch history [29]. VTAGE directly inherits the structure of the indirect jump predictor ITTAGE [10]. VTAGE is able to predict with a very high accuracy many values that were not correctly predicted by previously proposed predictors, such as the FCM predictor and the stride predictor. Three sources of information can be harnessed by these predictors: the global branch history, the differences of successive values and the local history of values. Moreover, VTAGE does not suffer from short critical prediction loops and can seamlessly handle back-to-back predictions, contrarily to previously proposed, hard to implement FCM predictors.

Second, we show that all predictors are amenable to very high accuracy at the cost of some loss on prediction coverage [29]. This greatly diminishes the number of value mispredictions and allows to delay validation until commit-time. As such, no complexity is added in the out-of-order engine because of VP (save for ports on the register file) and pipeline squashing at commit-time can be used to recover. This is crucial as adding *selective replay* in the OoO core would tremendously increase complexity.

Third, we leverage the possibility of validating predictions at commit to introduce a new microarchitecture, EOLE [28]. EOLE features *Early Execution* to execute simple instructions whose operands are ready in parallel with Rename and *Late Execution* to execute simple predicted instructions and high confidence branches just before Commit. EOLE depends on Value Prediction to provide operands for *Early Execution* and predicted instructions for *Late Execution*. However, Value Prediction requires EOLE to become truly practical. That is, EOLE allows to reduce the out-of-order issue-width by 33% without impeding performance. As such, the number of ports on the register file diminishes. Furthermore, optimizations of the register file such as *banking* further reduce the number of required ports. Overall EOLE possesses a register file whose complexity is on-par with that of a regular wider-issue superscalar while the out-of-order components (scheduler, bypass) are greatly simplified. Moreover, thanks to Value Prediction, speedup is obtained on many benchmarks of the SPEC'00/'06 suite.

6.2.1.3. Skewed Compressed Caches

Participant: André Seznec.

Cache compression seeks the benefits of a larger cache with the area and power of a smaller cache. Ideally, a compressed cache increases effective capacity by tightly compacting compressed blocks, has low tag and metadata overheads, and allows fast lookups. Previous compressed cache designs, however, fail to achieve all these goals. In this study, we propose the Skewed Compressed Cache (SCC), a new hardware compressed cache that lowers overheads and increases performance. SCC tracks super- blocks to reduce tag overhead, compacts blocks into a variable number of sub-blocks to reduce internal fragmentation, but retains a direct tag-data mapping to find blocks quickly and eliminate extra metadata (i.e., no backward pointers). SCC does this using novel sparse super-block tags and a skewed associative mapping that takes compressed size into account. In our experiments, SCC provides on average 8% (up to 22%) higher performance, and on average 6% (up to 20%) lower total energy, achieving the benefits of the recent Decoupled Compressed Cache [47] with a factor of 4 lower area overhead and lower design complexity.

This study was done in collaboration with Somayeh Sardashti and David Wood from University of Wisconsin.

6.2.1.4. Efficient Execution on Guarded Instruction Sets

Participant: André Seznec.

ARM ISA based processors are no longer low complexity processors. Nowadays, ARM ISA based processor manufacturers are struggling to implement medium-end to high-end processor cores which implies implementing a state-of-the-art out-of-order execution engine. Unfortunately providing efficient out-of-order execution on legacy ARM codes may be quite challenging due to guarded instructions.

Predicting the guarded instructions addresses the main serialization impact associated with guarded instructions execution and the multiple definition problem. Moreover, guard prediction allows to use a global branch-and-guard history predictor to predict both branches and guards, often improving branch prediction accuracy. Unfortunately such a global branch-and-guard history predictor requires the systematic use of guard predictions. In that case, poor guard prediction accuracy would lead to poor overall performance on some applications.

Building on top of recent advances in branch prediction and confidence estimation, we propose a hybrid branch and guard predictor, combining a global branch history component and global branch-and-guard history component. The potential gain or loss due to the systematic use of guard prediction is dynamically evaluated at run-time. Two computing modes are enabled: systematic guard prediction use and high confidence only guard prediction use. Our experiments show that on most applications, an overwhelming majority of guarded instructions are predicted. Therefore a relatively inefficient but simple hardware solution can be used to execute the few unpredicted guarded instructions. Significant performance benefits are observed on most applications while applications with poorly predictable guards do not suffer from performance loss [7].

This study is accepted to ACM Transactions on Architecture and Compiler Optimizations (to appear January 2015) and will be presented at the HIPEAC conference in January 2015.

6.2.1.5. Clustered microarchitecture

Participants: Andrea Mondelli, Pierre Michaud, André Seznec.

In the last 10 years, the clock frequency of high-end superscalar processors did not increase significantly. Performance keeps being increased mainly by integrating more cores on the same chip and by introducing new instruction set extensions. However, this benefits only to some applications and requires rewriting and/or recompiling these applications. A more general way to increase performance is to increase the IPC, the number of instructions executed per cycle.

We argue that some of the benefits of technology scaling should be used to increase the IPC of future superscalar cores. Starting from microarchitecture parameters similar to recent commercial high-end cores, we show that an effective way to increase the IPC is to increase the issue width. But this must be done without impacting the clock cycle. We propose to combine two known techniques: clustering and register write specialization. The objective of past work on clustered microarchitecture was to allow a higher clock frequency while minimizing the IPC loss. This led researchers to consider narrow-issue clusters. Our objective, instead, is to increase the IPC without impacting the clock cycle, which means wide-issue clusters. We show that, on a wide-issue dual cluster, a very simple steering policy that sends 64 consecutive instructions to the same cluster, the next 64 instructions to the other cluster, and so on, permits tolerating an inter-cluster delay of several cycles. We also propose a method for decreasing the energy cost of sending results of one cluster to the other cluster.

This work is currently under submission.

6.2.1.6. Adaptive Intelligent Memory Systems

Participants: André Seznec, Aswinkumar Sridharan.

On multicores, the processors are sharing the memory hierarchy, buses, caches, and memory. The performance of any single application is impacted by its environment and the behavior of the other applications co-running on the multicore. Different strategies have been proposed to isolate the behavior of the different co-running applications, for example performance isolation cache partitioning, while several studies have addressed the global issue of optimizing throughput through the cache management.

However these studies are limited to a few cores (2-4-8) and generally feature mechanisms that cannot scale to 50-100 cores. Moreover so far the academic propositions have generally taken into account a single parameter, the cache replacement policy or the cache partitioning. Other parameters such as cache prefetching and its aggressiveness already impact the behavior of a single thread application on a uniprocessor. Cache prefetching policy of each thread will also impact the behavior of all the co-running threads.

Our objective is to define an Adaptive and Intelligent Memory System management hardware, AIMS. The goal of AIMS will be to dynamically adapt the different parameters of the memory hierarchy access for each individual co-running process in order to achieve a global objective such as optimized throughput, thread fairness or respecting quality of services for some privileged threads.

6.2.2. Microarchitecture Performance Modeling

6.2.2.1. Multiprogram throughput of multicore/SMT processors

Participant: Pierre Michaud.

This research was done in collaboration with Stijn Eyerman and Wouter Rogiest from Ghent University.

There are several aspects to the performance of a multicore processor. One of them is multiprogram throughput, that is, how fast a multicore can execute several independent jobs. However, defining throughput metrics that are both meaningful and practical for computer architecture studies is not straightforward. We present a method to construct throughput metrics in a systematic way: we start by expressing assumptions on job sizes, job types distribution, scheduling, etc., that together define a theoretical throughput experiment. The throughput metric is then the average throughput of this experiment. Different assumptions lead to different metrics, so one should be aware of these assumptions when making conclusions based on results using a specific metric. Throughput metrics should always be defined from explicit assumptions, because this leads to a better understanding of the implications and limits of the results obtained with that metric. We elaborate multiple metrics based on different assumptions. In particular, we show that commonly used throughput metrics such as instructions per cycle and weighted speedup implicitly assume a variable workload, that is, a workload which depends on the machine being evaluated. However, in many situations, it is more realistic to assume a fixed workload. Hence we propose some new fixed-workload throughput metrics. Evaluating these new metrics requires to solve a continuous-time Markov chain. We released a software, TPCalc, that takes as input the performance results of individual coschedules simulations and computes fixed-workload throughput, taking advantage of multicore symmetries [15].

In a subsequent work, we applied our framework to symbiotic scheduling on a symmetric multicore or SMT processor. Symbiotic scheduling tries to exploit the fact, because of resource sharing (execution units, caches, memory bandwidth, etc.) and because different jobs have different characteristics, the performance may be increased by carefully choosing the coschedules. We show that, when assuming a fixed workload, an optimal schedule maximizing throughput can be found by solving a linear programming problem. However, the throughput gains we observed in our experiments, 3% on average, are significantly smaller than what we expected based on published studies on symbiotic scheduling. We analyzed the reasons for this and we found the two main reasons for this discrepancy: previous studies either did not consider a fixed workload but a variable one, or did not report throughput gains but response time reductions. Response time reductions can be artificially magnified by setting the job arrival rate close to the maximum throughput.

This work will be presented at the ISPASS 2015 conference.

6.2.2.2. Modeling multi-threaded programs execution time in the many-core era

Participants: Surya Khizakanchery Natarajan, Bharath Narasimha Swamy, André Seznec.

Estimating the potential performance of parallel applications on the yet-to-be-designed future many cores is very speculative. The traditional laws used to predict performance of an application do not reflect on the various scaling behaviour of a multi-threaded (MT) application leading to optimistic estimation of performance in manycore era. In this paper, we study the scaling behavior of MT applications as a function of input workload size and the number of cores. For some MT applications in the benchmark suites we analysed, our study shows that the serial fraction in the program increases with input workload size and can be a scalability-limiting factor. Similar to previous studies [41] , we find that using a powerful core (heterogeneous architecture) to execute this serial part of the program can mitigate the impact of serial scaling and improve the overall performance of an application in many-core era [25].

6.2.3. Hardware/Software Approaches

6.2.3.1. Helper threads

Participants: Bharath Narasimha Swamy, Alain Ketterlin, André Seznec.

Heterogeneous Many Cores (HMC) architectures that mix many simple/small cores with a few complex/large cores are emerging as a design alternative that can provide both fast sequential performance for single threaded workloads and power-efficient execution for throughput oriented parallel workloads. The availability of many small cores in a HMC presents an opportunity to utilize them as low-power helper cores to accelerate memory-intensive sequential programs mapped to a large core. However, the latency overhead of accessing small cores in a loosely coupled system limits their utility as helper cores. Also, it is not clear if small cores can execute helper threads sufficiently in advance to benefit applications running on a larger, much powerful, core. In [24], we present a hardware/software framework called core-tethering to support efficient helper threading on heterogeneous many-cores. Core-tethering provides a co-processor like interface to the small cores that (a) enables a large core to directly initiate and control helper execution on the helper core and (b) allows efficient transfer of execution context between the cores, thereby reducing the performance overhead of accessing small cores for helper execution. Our evaluation on a set of memory intensive programs chosen from the standard benchmark suites show that, helper threads using moderately sized small cores can significantly accelerate a larger core compared to using a hardware prefetcher alone. We find that a small core provides a good trade-off against using an equivalent large core to run helper threads in a HMC. Additionally, helper prefetching on small cores when used along with hardware prefetching, can provide an alternate design point to growing instruction window size for achieving higher sequential performance on memory intensive applications.

6.2.3.2. Branch Prediction and Performance of Interpreter

Participants: Erven Rohou, André Seznec, Bharath Narasimha Swamy.

Interpreters have been used in many contexts. They provide portability and ease of development at the expense of performance. The literature of the past decade covers analysis of why interpreters are slow, and many software techniques to improve them. A large proportion of these works focuses on the dispatch loop, and in particular on the implementation of the switch statement: typically an indirect branch instruction. Folklore attributes a significant penalty to this branch, due to its high misprediction rate. We revisit this assumption, considering state-of-the-art branch predictors and the three most recent Intel processor generations on current interpreters. Using both hardware counters on Haswell, the latest Intel processor generation, and simulation of the ITTAGE, we show that the accuracy of indirect branch prediction is no longer critical for interpreters. We further compare the characteristics of these interpreters and analyze why the indirect branch is less important than before.

This study [8] has been accepted for publication at CGO 2015 (International Symposium on Code Generation and Optimization).

6.2.3.3. Augmenting superscalar architecture for efficient many-thread parallel execution

Participants: Sylvain Collange, André Seznec, Sajith Kalathingal.

We aim at exploring the design of a unique core that efficiently runs both sequential and massively parallel sections. We explore how the architecture of a complex superscalar core has to be modified or enhanced to efficiently run several threads from the same application.

Rather than vectorize at compile-time, our approach is to dynamically vectorize SPMD programs at the micro-architectural level. The SMT-SIMD hybrid core we propose extracts data parallelism from thread parallelism by scheduling groups of threads in lockstep, in a way inspired by the execution model of GPUs. As in GPUs, conditional branches whose outcomes differ between threads are handled with conditionally masked execution. However, while GPUs rely on explicit re-convergence instructions to restore lockstep execution, we target existing general-purpose instruction sets, in order to run legacy binary programs. Thus, the main challenge consists in detecting re-convergence points dynamically.

To handle this difficulty, we can build on [17]. In this work done in collaboration with Fernando Pereira and his team at UFMG, Brasil, we proposed instruction fetch policies that apply heuristics to maximize the cycles spent in lockstep execution, and evaluated them under a micro-architecture independent model [17]. Results highlight the necessity of a tradeoff between maximizing throughput and extracting data-level parallelism with lockstep execution.

6.3. Compiler, vectorization, interpretation

Participants: Erven Rohou, Emmanuel Riou, Bharath Narasimha Swamy, Arjun Suresh, André Seznec, Nabil Hallou, Alain Ketterlin, Sylvain Collange.

6.3.1. Compilers for emerging throughput architectures

Participant: Sylvain Collange.

This work is done in collaboration with Fernando Pereira and his team at UFMG, Brasil.

GPU architectures present new challenges for compilers. Their performance characteristics demand SPMD programs with a high control-flow and memory regularity. Such architecture takes advantage of the regularity in programs to exploit data-level parallelism. In addition to the traditional challenges of code parallelization, new compilers for GPU and future throughput architectures face the task of improving the regularity of parallel programs. In particular, compiler analyses that identify control-flow divergence and memory divergence are a stepping stone for many optimizations. These optimizations include traditional code transformation such as loop interchange and tiling, which use divergence as an additional decision criterion, but also new optimizations specific to GPU architectures such as iteration delaying or branch fusion. In addition, the regularity parameter is an important aspect for workload characterization, as it provides a criterion for task scheduling in heterogeneous environments, such as multi-core processors with GPU. Our objectives include both accurate static and dynamic analyses for thread divergence, and the applications that it enables. We propose to combine static analyses with runtime checks, in order to get the best from both complementary approaches.

6.3.2. Improving sequential performance through memoization

Participants: Erven Rohou, André Seznec, Arjun Suresh.

Many applications perform repetitive computations, even when properly programmed and optimized. Performance can be improved by caching results of pure functions, and retrieving them instead of recomputing a result (a technique called memoization).

We proposed a simple technique for enabling software memoization of any dynamically linked pure function and we illustrate our framework using a set of computationally expensive pure functions – the transcendental functions.

Our technique does not need the availability of source code and thus can be applied even to commercial applications as well as applications with legacy codes. As far as users are concerned, enabling memoization is as simple as setting an environment variable.

Our framework does not make any specific assumptions about the underlying architecture or compiler tool-chains, and can work with a variety of current architectures.

We present experimental results for x86-64 platform using both gcc and icc compiler tool-chains, and for ARM cortex-A9 platform using gcc. Our experiments include a mix of real world programs and standard benchmark suites: SPEC and Splash2x. On standard benchmark applications that extensively call the transcendental functions we report memoization benefits of upto 16 %, while much higher gains were realized for programs that call the expensive Bessel functions. Memoization was also able to regain a performance loss of 76 % in *bwaves* due to a known performance bug in the gcc libm implementation of *pow* function.

6.3.3. Code Obfuscation

Participant: Erven Rohou.

This research is done in collaboration with the group of Prof. Ahmed El-Mahdy at E-JUST, Alexandria, Egypt.

A new obfuscation technique [27] based of decomposition of CFGs into threads has been proposed. We exploit the mainstream multi-core processing in these systems to substantially increase the complexity of programs, making reverse engineering more complicated. The novel method automatically partitions any serial thread into an arbitrary number of parallel threads, at the basic-block level. The method generates new control-flow graphs, preserving the blocks' serial successor relations and guaranteeing that one basic-block is active at a time through using guards. The method generates m^n different combinations for m threads and n basic-blocks, significantly complicating the execution state. We also provide proof of correctness for the method.

We propose to leverage JIT compilation to make software tamper-proof. The idea is to constantly generate different versions of an application, even while it runs, to make reverse engineering hopeless. More precisely a JIT engine is used to generate new versions of a function each time it is invoked, applying different optimizations, heuristics and parameters to generate diverse binary code. A strong random number generator will guarantee that generated code is not reproducible, though the functionality is the same.

This work has been accepted for publication in January 2015 at the International Workshop on Dynamic Compilation Everywhere (DCE-2015).

6.3.4. Padrone

Participants: Erven Rohou, Alain Ketterlin, Emmanuel Riou.

The objective of the ADT PADRONE is to design and develop a platform for re-optimization of binary executables at run-time. Development is ongoing, and an early prototype is functional. In [30], we described the infrastructure of Padrone, and showed that its profiling overhead is minimum. We illustrated its use through two examples. The first example shows how a user can easily write a tool to identify hotspots in their application, and how well they perform (for example, by computing the number of executed instructions per cycle). In the second example, we illustrate the replacement of a given function (typically a hotspot) by an optimized version, while the program runs.

We believe PADRONE fills an empty design point in the ecosystem of dynamic binary tools.

6.3.5. Dynamic Binary Re-vectorization

Participants: Erven Rohou, Nabil Hallou, Alain Ketterlin, Emmanuel Riou.

This work is done in collaboration with Philippe Clauss (Inria CAMUS).

Applications are often under-optimized for the hardware on which they run. Several reasons contribute to this unsatisfying situation, including the use of legacy code, commercial code distributed in binary form, or deployment on compute farms. In fact, backward compatibility of instruction sets guarantees only the functionality, not the best exploitation of the hardware. In particular SIMD instruction sets are always evolving.

We proposed a runtime re-vectorization platform that dynamically adapts applications to execution hardware. Programs distributed in binary forms are re-vectorized at runtime for the underlying execution hardware. Focusing on the x86 SIMD extensions, we are able to automatically convert loops vectorized for SSE into the more recent and powerful AVX. A lightweight mechanism leverages the sophisticated technology put in a static vectorizer and adjusts, at minimal cost, the width of vectorized loops. We achieve speedups in line with a native compiler targeting AVX. Our re-vectorizer is implemented inside a dynamic optimization platform; its usage is completely transparent to the user and requires neither access to source code nor rewriting binaries.

6.4. WCET estimation

Participants: Damien Hardy, Hanbing Li, Isabelle Puaut, Erven Rohou.

Predicting the amount of resources required by embedded software is of prime importance for verifying that the system will fulfill its real-time and resource constraints. A particularly important point in hard real-time embedded systems is to predict the Worst-Case Execution Times (WCETs) of tasks, so that it can be proven that tasks temporal constraints (typically, deadlines) will be met. Our research concerns methods for obtaining automatically upper bounds of the execution times of applications on a given hardware. Our new results this year are on (i) multi-core architectures (ii) WCET estimation for faulty architectures (iii) traceability of flow information in compilers for WCET estimation.

6.4.1. WCET estimation and its interactions with compilation

6.4.1.1. On the comparison of deterministic and probabilistic WCET estimation techniques

Participants: Damien Hardy, Isabelle Puaut.

This is joint work with Jaume Abella, Eduardo Quinones and Francisco J. Cazorla from Barcelona Supercomputing Center

Several timing analysis techniques have been proposed to obtain Worst-Case Execution Time (WCET) estimates of applications running on a particular hardware. They can be classified into two classes of approaches: deterministic timing analysis techniques (DTA), that produce a unique WCET estimate, and probabilistic timing analysis techniques (PTA) that produce multiple WCET estimates with associated probabilities. Both approaches have their static (SDTA, SPTA) and measurement-based (MBDTA, MBPTA) variants. The lack of comparison figures among those techniques makes complex the selection of the most appropriate one.

This work [19] makes a first attempt towards comparing comprehensively SDTA, SPTA and MBPTA qualitatively and quantitatively, under different cache configurations implementing LRU and random replacement. We identify strengths and limitations of each technique depending on the characteristics of the program under analysis and the hardware platform, thus providing users with guidance on which approach to choose depending on their target application and hardware platform.

6.4.2. WCET estimation for architectures with faulty caches

Participants: Damien Hardy, Isabelle Puaut.

Technology scaling, used to increase performance, has the negative consequence of providing less reliable silicon primitives, resulting in an increase of the probability of failure of circuits, in particular for SRAM cells. While space redundancy techniques exist to recover from failures and provide fault-free chips, they will not be affordable anymore in the future due to their growing cost. Consequently, other approaches like fine grain disabling and reconfiguration of hardware elements (e.g. individual functional units or cache blocks) will become economically necessary. This fine-grain disabling will lead to degraded performance compared to a fault-free execution.

A common implicit assumption in all static worst-case execution time (WCET) estimation methods is that the target processor is not subject to faults. Their result is not safe anymore when using fine grain disabling of hardware components, which degrades performance.

In [16] a method that statically calculates a probabilistic WCET bound in the presence of permanent faults in instruction caches is provided. The method, from a given program, cache configuration and probability of cell failure, derives a probabilistic WCET bound. An essential benefit of our approach is that its probabilistic nature stems only from the probability associated with the presence of faults. By construction, the worst-case execution path cannot be missed, since it is determined using static analysis, extended to cope with permanent faults. This allows our method to be used in safety-critical real-time systems. The method is computationally tractable, since it avoids the exhaustive enumeration of all possible fault locations. Experimental results show that the proposed method accurately estimates WCETs in the presence of permanent faults compared to a method that explores all possible locations for faults. On the one hand, the proposed method allows to quantify the impact of permanent faults on WCET estimates for chips with a known probability of cell failure for the whole chip lifetime. On the other hand, and most importantly, our work can also be used in architectural exploration frameworks to select the most appropriate fault management mechanisms, for current and future chip designs.

6.4.3. Traceability of flow information for WCET estimation

Participants: Hanbing Li, Isabelle Puaut, Erven Rohou.

This research is part of the ANR W-SEPT project.

Control-flow information is mandatory for WCET estimation, to guarantee that programs terminate (e.g. provision of bounds for the number of loop iterations) but also to obtain tight estimates (e.g. identification of infeasible or mutually exclusive paths). Such flow information is expressed though annotations, that may be calculated automatically by program/model analysis, or provided manually.

The objective of this work is to address the challenging issue of the mapping and transformation of the flow information from high level down to machine code. In our recent work [21], we have proposed a framework to systematically transform flow information from source code to machine code.

The framework defines a set of formulas to transform flow information for standard compiler optimizations. Transforming the flow information is done within the compiler, in parallel with transforming the code. There thus is no guessing what flow information have become, it is transformed along with the code. The framework is general enough to cover all linear flow constraints and all typical optimizations implemented in modern compilers. Our implementation in the LLVM compiler shows that we can improve the WCET of Malardalen benchmarks by 60% in average (up to 86%) by turning on optimizations. We also provide new insight on the impact of existing optimizations on the WCET.

6.4.4. Verified WCET estimation

Participant: Isabelle Puaut.

This is joint work with Andre Oliveira Maroneze, David Pichardie and Sandrine Blazy from the Celtique group at Inria/IRISA Rennes.

Current WCET estimation tools, even when based on sound static analysis techniques, are not verified. This may lead to bugs being accidentally introduced in the implementation. The main contribution of this work [13], [26] is a formally verified WCET estimation tool operating over C code.

Our tool is integrated to the formally verified CompCert C compiler. It is composed of two main parts: a loop bound estimation and an Implicit Path Enumeration Technique (IPET)-based WCET calculation method. We evaluated the precision of the WCET estimates on a reference benchmark and obtained results which are competitive with state-of-the-art WCET estimation techniques. The code of our tool is automatically generated from its formal specification. Furthermore, machine-checked proofs ensure the estimated WCET is at least as large as the actual WCET.

6.5. Computer arithmetic

Participant: Sylvain Collange.

6.5.1. Application-specific number systems

Collaboration with Mark G. Arnold, XLNS Research, USA.

Reconfigurable FPGA platforms let designers build efficient application-specific circuits, when the performance or energy efficiency of general-purpose CPUs is insufficient, and the production volume is not enough to offset the very high cost of building a dedicated integrated circuit (ASIC). One way to take advantage of the flexibility offered by FPGAs is to tailor arithmetic operators for the application. In particular, the Logarithmic Number System (LNS) is suitable for embedded applications dealing with low-precision, high-dynamic range numbers.

Like floating-point, LNS can represent numbers from a wide dynamic range with constant relative accuracy. However, while standard floating-point offer so-called subnormal numbers to represent numbers close to zero with constant absolute accuracy, LNS numbers abruptly overflow to zero, resulting in a gap in representable numbers close to zero that can impact the accuracy of numerical algorithms.

In collaboration with Mark G. Arnold, Sylvain Collange proposed a generalization of LNS that incorporates features analogous to subnormal floating-point [14]. The Denormal LNS (DLNS) system we introduce defines a class of hybrid number systems that offer quasi-constant absolute accuracy close to zero and quasi-constant relative accuracy on larger numbers. These systems can be configured to range from pure LNS (constant relative accuracy) to fixed-point (constant absolute accuracy across the whole range).

6.5.2. Deterministic floating-point primitives for high-performance computing

Parallel algorithms such as reduction are ubiquitous in parallel programming, and especially high-performance computing. Although these algorithms rely on associativity, they are use on floating-point data, on which operations are not associative. As a result, computations become non-deterministic, and the result may change according to static and dynamic parameters such as machine configuration or task scheduling.

In collaboration with David Defour (UPVD), Stef Graillat and Roman Iakymchuk (LIP6), we introduce a solution to compute deterministic sums of floating-point numbers efficiently and with the best possible accuracy. A multi-level algorithm incorporating a filtering stage that uses fast vectorized floating-point expansions and an accumulation stage based on super-accumulators in a high-radix carry-save representation guarantees accuracy to the last bit even on degenerate cases while maintaining high performance in the common cases [35].

ALGORILLE Project-Team

6. New Results

6.1. Structuring applications for scalability

6.1.1. Combining locking and data management interfaces

Participants: Jens Gustedt, Mariem Saied.

Handling data consistency in parallel and distributed settings is a challenging task, in particular if we want to allow for an easy to handle asynchronism between tasks. Our publication [4] shows how to produce deadlock-free iterative programs that implement strong overlapping between communication, IO and computation.

A new implementation (ORWL) of our ideas of combining control and data management in C has been undertaken, see 5.2.1 . In 2014, work has demonstrated its efficiency for a large variety of platforms, see [20]. By using the example of dense matrix multiplication, we show that ORWL permits to reuse existing code for the target architecture, namely open source library ATLAS, Intel's compiler specific MKL library or NVidia's CUBLAS library for GPUs. ORWL assembles local calls into these libraries into efficient functional code, that combines computation on distributed nodes with efficient multi-core and accelerator parallelism.

Our next efforts will concentrate on the continuation of an implementation of a complete application (an American Option Pricer) that was chosen because it presents a non-trivial data transfer and control between different compute nodes and their GPU. ORWL is able to handle such an application seamlessly and efficiently, a real alternative to home made interactions between MPI and CUDA.

6.2. Experimental methodologies for the evaluation of distributed systems

6.2.1. Simulation and dynamic verification

6.2.1.1. SimGrid framework improvement

Participants: Paul Bédaride, Martin Quinson, Gabriel Corona.

On the technical side, we kept up with our regular releases of the SimGrid framework, integrating the work of our partners in the SONGS ANR project. This year, we reimplemented the simulation kernel in C++. This modularity improvement will ease the addition of performance models by external contributors. This work thus contributes to our overall goal of constituting a user community focused on this first-class tool.

[11] is a long awaited paper describing the current state of the project and its future roadmap. This constitutes the new reference paper on the SimGrid project (the previous article, a short paper from 2008, was cited over 350 times since its publication). We show that despite the common beliefs, the tool specialization is not necessarily a warrant for performance and correctness.

We also continued our animation of our scientific community, for example through our participation to the Joint Laboratory for Petascale Computing (Inria/ANL/UIUC/BSC). We co-organized a summer school on Performance Metrics, Modeling and Simulation of Large HPC Systems in June, to push our tools toward PhD students that need to assess their HPC applications.

6.2.1.2. Dynamic verification and SimGrid

Participants: Marion Guthmuller, Martin Quinson, Gabriel Corona.

This year, the PhD thesis of M. Guthmuller went into its third year. The proposed methodology maturated into a usable tool: we can now verify small-size real HPC applications using MPI in C/C++/Fortran. This relies on a heuristic exploration of the applicative state at the system level that was presented in [21], [22].

Also, we finally added the ability to dynamically verify some CTL properties over MPI implementations. SimGrid was one of the rare framework able to verify LTL liveness properties over real implementations. To the best of our knowledge, it becomes the very first tool verifying CTL properties on real C/C++/Fortran applications. The targeted properties quantify the stability of the applicative communication pattern. The applications that respect these properties can benefit from specific, more efficient, fault tolerance algorithms. Verifying these properties is thus of a major practical interest. A publication is in preparation, as well as the PhD manuscript of M. Guthmuller who will defend by 2015 Q1.

6.2.2. Experimentation on testbeds and production facilities, emulation

6.2.2.1. Evaluating load balancing and fault tolerance strategies on Distem

Participants: Joseph Emeras, Emmanuel Jeanvoine, Lucas Nussbaum.

(For context, see sections 3.3 and 5.4 .)

We extended our work [27] to enable the study of load balancing and fault tolerance strategies on Distem. Distem now supports the introduction of changing heterogeneity and imbalance among virtual nodes, as well as the introduction of failures. Two HPC runtimes targeting Exascale (Charm++ and OpenMPI) were used as target applications. This work was presented at the Joint Laboratory for Extreme-Scale Computing in June, and at the Grid'5000 Spring School. However, those results still have to be properly published.

6.2.2.2. Distem improvements: VXLAN, release and tutorial

Participants: Emmanuel Jeanvoine, Tomasz Buchert, Lucas Nussbaum.

(For context, see sections 3.3 and 5.4 .)

The scalability of Distem's networking layer was improved by adding support for VXLAN networks. This enabled experiments with up to 40,000 virtual nodes, presented at the CCGrid'2014 SCALE challenge (where we were selected as finalist) [17]. Version 1.0 of Distem was also released in March 2014, and featured in a tutorial at the Grid'5000 Spring School.

6.2.2.3. Kadeploy improvements: REST API, new image broadcast mechanism

Participants: Luc Sarzyniec, Stéphane Martin, Emmanuel Jeanvoine, Lucas Nussbaum.

(For context, see sections 3.3 and 5.4 .)

Kadeploy 3.2 was released in March 2014. Among many other changes, that release included a new REST API to interact with Kadeploy, replacing the old Ruby-specific RPC mechanism, and easing the automation of experiments by providing a way to call Kadeploy from scripts.

Kadeploy 3.3 was released in November 2014. This release is mostly a bug-fix release, with many bug fixes in the internal cache system, the shell runner, and others.

We also implemented an improved mechanism to broadcast machine images to nodes. The new tool, called Kascade, is fault tolerant, and its performance has been thoroughly tested. It was described in a publication accepted at HPDIC'2014 [24], included in Kadeploy 3.2, and used as the default method for environment broadcast since Kadeploy 3.3.

6.2.2.4. XPFlow

Participants: Tomasz Buchert, Stéphane Martin, Emmanuel Jeanvoine, Lucas Nussbaum, Jens Gustedt.

(For context, see sections 3.3 and 5.7 .)

A publication focusing on XPFlow was accepted at CCGrid'2014 [18], and XPFlow was also featured in a tutorial at Grid'5000 Spring School. Our ongoing work focuses on improved support for collecting provenance in XPFlow.

6.2.2.5. Survey of Experiment Management tools

Participants: Tomasz Buchert, Cristian Ruiz, Lucas Nussbaum.

We produced a survey of Experiment Management tools for distributed systems, published in Future Generation Computer Systems [10]. This survey provides an extensive list of features offered by general-purpose experiment management tools dedicated to distributed systems research on real platforms. It then uses it to assess existing solutions and compare them, outlining possible future paths for improvements.

6.2.2.6. Grid'5000

Participants: Émile Morel, Luc Sarzyniec, Lucas Nussbaum.

(For context, see sections 3.3 and 5.8 .)

The work on resources description, selection, reservation and verification was wrapped-up in a TridentCom'2014 paper [23].

As a member of the Grid'5000 architects committee, Lucas Nussbaum was involved in the submission (and acceptance) of ADT Laplace.

Lucas Nussbaum also presented a talk [12] on Reproducible Research and Grid'5000 at the Grid'5000 evaluation by the Scientific Committee, during the Spring School.

6.2.3. Convergence and co-design of experimental methodologies

6.2.3.1. Realis'2014

Participant: Lucas Nussbaum.

Lucas Nussbaum organized (with Olivier Richard) the second edition of the Realis event [14]. Associated to the Compas'14 conference, this workshop aimed at providing a place to discuss the reproducibility of the experiments underlying the publications submitted to the main conference. We hope that this kind of venue will motivate the researchers to further detail their experimental methodology, ultimately allowing others to reproduce their experiments.

6.2.3.2. Reproducible Research working group at Inria Nancy – Grand Est

Participant: Lucas Nussbaum.

Lucas Nussbaum is organizing a working group on Reproducible Research at Inria Nancy – Grand Est since May 2014. Meetings involve a dozen of members from many different teams, and discussion topics have so far covered online platforms to test algorithms and applications, and evaluation contests organized together with conferences and workshops.

Lucas Nussbaum has also been invited to participate in the Inria national initiative on reproducible research.

6.2.3.3. Organization of Reppar

Participant: Lucas Nussbaum.

Lucas Nussbaum co-organized the first edition of the Reppar workshop, held during Europar'2014, with a focus on experimental practices in parallel computing research.

6.3. Algorithmic schemes for efficient use of parallel devices in clusters

Participants: Sylvain Contassot-Vivier, Stéphane Vialle [External collaborator, SUPELEC].

During the year 2014, we have continued our studies about the design and implementation of efficient algorithmic schemes to fully exploit all the available computational resources inside a parallel system. In particular, we have proposed general schemes that optimize the use of GPUs in clusters [26]. This is achieved by performing two kinds of overlappings. The former corresponds to computation/communication overlappings, either for the communications between machines but also for the data transfers between central RAM and GPUs inside each machine. The latter is the computation/computation overlapping that consists in executing computations on the GPUs in parallel of some computations on the central CPUs. Moreover, in this work we have paid a particular attention to some important aspects of software engineering that are the development and maintenance costs. Those aspects are essential as they directly determine the practical usability of the schemes, especially in the industry where there is a permanent vigilance to minimize the associated costs.

6.4. Parallel schemes for the resolution of the RTE with finite volumes method

Participant: Sylvain Contassot-Vivier.

In the context of our collaboration with the Lemta laboratory (Fatmir Asllanaj), about the design and implementation of an efficient and high accuracy algorithm for solving the Radiative Transfer Equation (RTE), we have reached our second objective that consisted in the realization of a multi-threaded parallel version of the software. That new version is based on the optimized sequential version produced as a first objective. It makes use of the OpenMP library to exploit all the cores inside one machine. The results are very satisfying as our algorithm obtains very good speed up and efficiency (around 90% and above) in realistic contexts. Moreover, besides this work over performance, we focus also on the high quality (accuracy) of the results of our software by making a permanent effort to track any possible enhancement of our numerical scheme. Then, the actual implementation of each of these possible enhancements is considered according to its potential costs, either in performance degradation as well as in additional resource consumptions (CPUs, GPUs and RAM). Confrontations to other existing computational schemes to solve the RTE are regularly realized to corroborate the validity preservation of our software [9], [15].

6.5. Study of binary multiplication and dynamical approaches to the integer factorization

Participants: Sylvain Contassot-Vivier, Nazim Fatès.

In the context of a collaboration with Nazim Fatès over dynamical systems we have co-supervised the internship of Raphaël Rieu-Helft (student at the ENS Paris), during June and July 2014. The goal of this internship was to study the relevance of the dynamical systems formalism as an efficient way to express and solve two specific problems. The former one was the queens problem on chessboards of arbitrary size. This goal was to express a solving algorithm of the queens problem under the form of a cellular automaton. The second step was to extend the results obtained for the queens problem to a more complex and computationally expensive problem that is the integer factorization. Two dynamical systems (cellular automata) have been obtained for both problems and their respective efficiencies, either in terms of convergence speed or speed of solution reaching, have been experimentally evaluated.

ALICE Project-Team

6. New Results

6.1. Highlights of the Year

Fabrication: We proposed a novel technique to automatically generate support structures for additive manufacturing with filament based processes. The deposited filament has to be properly supported at all times, which complicates printing of overhanging shapes: a disposable support has to be generated to temporarily hold the filament deposited above. Existing techniques either generate large structures, wasting material, or generate very thin structures that are hard to print and prone to failure. In contrast, our technique optimizes a scaffolding which is made of vertical pillars and horizontal bridges – such horizontal bridges print properly as long and the filament is deposited in straight line from one pilar to the next. We showed how to formulate scaffolding generation as a minimization problem and proposed a heuristic algorithm based on an efficient plane sweeping approach. The work was published [9] in ACM Transactions on Graphics in 2014 (proceedings of SIGGRAPH 2014). It is integrated within our 3D modeler for additive manufacturing, IceSL.

Optimal transport: this is an active research topics in the mathematics community. Given two measures μ and ν, optimal transport defines a distance between μ and ν, as the minimum cost of “morphing” μ into ν. This distance (called the *Wasserstein distance*) structures the space of measures and offers new ways of solving some highly non-linear PDEs (Monge-Ampere, Fokker-Plank ...). This requires a numerical way of computing the Wasserstein distance and its gradients. We studied a semi-discrete technique [21] submitted to ESAIM J. M2AN), that optimizes power diagrams. This is to our knowledge the first numerical implementation of optimal transport for volumetric densities (computes the Wasserstein distance between a sum of Dirac masses and a piece-wise linear density supported on a tetrahedral mesh).

6.2. New results

This year, we obtained new results in fabrication, in geometry processing and in multi-view reconstruction.

We investigated software solutions for printing with low cost (filament) 3D printers. We proposed a solution to automatically define temporary structures that will supports the object during its creation [9]. We also strongly reduce the artefacts that are produced by multi-material printing [17]. These works allow to better understand the physics of these printers, and to come up with efficient software solutions to common drawbacks of this technology. Other contributions in fabrication are more related to the design of the printable objects, that is developed in our software IceSL. To achieve real-time rendering of CSG models, we developed a new GPU approach for single pass A-Buffer [23]. This technique is also a simple solution to handle complex rendering problems such as transparency. We also proposed [11] an efficient method for performing dilatation and erosion directly on the same representation of volume by sequence of dilatation and erosions on segments.

In geometry processing, we proposed an algorithm to compute the intersection of Voronoi cells and a simplicial complex [25]. This algorithm is fast in dimension up to $10D$ because it doesn’t require to explicit the Voronoi diagram. It comes with exact predicates and symbolic perturbation to ensure its robustness. We have also developed an algorithm [13] able to trace streamlines on triangulated surfaces in such a way that two such streamlines cannot cross or merge. This property seems obvious in the continuous case, but was very difficult to enforce with the discrete representations (triangulated surface, and floating points) manipulated by the computer. We did also revisit the Optimal Delaunay triangulation in the case of graded mesh generation [14], and we adapted our remeshing methods to Geologic applications [27].

We obtained some new results in multi-view reconstruction: a new method that expands a limited set of correspondences towards a quasi-dense map across two views [15], and an improvement of variational multi-view reconstruction obtained thanks to a simple characterization of geometric deformations [16].

ALPAGE Project-Team

6. New Results

6.1. Highlights of the Year

Benoit Crabbé is a Junior Member of the Institut Universitaire de France (IUF) since October 2014. Two out of the five academic staff at Alpage are now member of the IUF, Laurence Danlos being a Senior Member since October 2013.

6.2. Automatic text normalisation

Participants: Benoît Sagot, Marion Baranes.

Since the emergence of the web, one of the goals of natural language processing (NLP) tools has been analysing raw noisy text documents such as blogs, review sites or social networks. These texts commonly contain misspellings, redundant punctuation, smileys, etc. Consequently they require specific preprocessing before being used in different NLP applications. That is why, we worked at Alpage on the development of a new corpora and the implementation of an automatic system for normalisation of such texts:

- **Corpus crap** In 2014, a large-scale extension of the number of normalisation rules used by the MElt part-of-speech tagger for processing noisy computer-generated content has been achieved. This work was carried out in the context of and based on corpora developed within the CoMeRe project, funded by the Institut de Linguistique Française and lead by Thierry Chanier [14].
- **Normalisation system** We have implemented a modular system which follows SxPipe [109]. This system detects if an unknown word to a reference lexicon corresponds to a non-word error (and is not a neologisme or a borrowing). Then, it attempts to normalize non-word errors and grammatical errors. In 2014, we focused on these two latter tasks. First, we have implemented a system which suggests one or several normalization candidates for these non-word errors. As described in [17], to do that, we use an analogy-based approach for acquiring normalisation rules and use them in the same way as lexical spelling correction rules. Secondly, we propose to normalize grammatical errors. To do that, we check for each word if it has common homophones. If this is the case, we consider these homophones as possible candidates for normalization. Finally, we filter all these candidates in order to keep only the one which is the most probable. This filtration is done using a probabilistic model based on a n-gram system. Moreover, the implementation of this system of normalisation motivated a side task. We developed an unsupervised method for acquiring pairs of lexical entries belonging to the same morphological family, i.e., derivationally related words, starting from a purely inflectional lexicon. This work, detailed in [16], allows us to create new linguistic resources for English, French, German and Spanish which contains derivational relations.

6.3. The impact of morphosyntactic processing on post-OCR error correction

Participants: Kata Gábor, Benoît Sagot, Pierre Magistry.

State of the art optical character recognition (OCR) software currently achieve an error rate of around 1 to 10% depending on the age and the layout of the text. To our knowledge, very little work has been done to exploit linguistic analysis for post-OCR error correction. Within the PACTE project we are conducting research on reducing the OCR error rate by using contextual information and linguistic processing.

In 2014 we continued our investigations on how named entity recognition can benefit OCR error detection by applying context-aware error correction rules directly to the OCR output. Several grammars have been created or improved to adress OCR problems occurring within different types of named entities. As a result, the SxPipe-PACTE toolchain was created to correct named entities in a noisy input [45], [31].

While the symbolic error correction method works with a very high precision, its limitation lies in its relatively low coverage. In order to deal with the errors occurring outside the recognized entitites, we studied the possibility of using lattice-based part of speech tagging to select the best correction hypothesis in context. Different methods were investigated to generate correction hypotheses, using word alignment software or by observing frequently occurring error types. The initial results confirm that a significant number of the remaining OCR errors can be corrected via lattice-based tagging, as long as the noise introduced by correction hypotheses is controlled.

6.4. Linear-time discriminant syntactico-semantic parsing

Participants: Benoit Crabbé, Maximin Coavoux, Djamé Seddah.

In this module we study efficient and accurate models of statistical phrase structure parsing. We focus on linear time lexicalized parsing algorithms (shift reduce, left corner) with approximations entailing linear time processing. The existing prototype involves a global discriminant parsing model of the large margin family (Perceptron,Mira, SVM avatars) able to parse user defined structured input tokens [23]. Thus the model can take into account various sources of information for taking decisions such as word form, part of speech, morphology or semantic classes inter alia.

Our participation to the SPRML 2014 shared task on parsing morphologically rich languages has been a first step towards testing our model in a multilingual setting where we were among the state of the art systems and state of the art on some languages such as Polish. To our knowledge the parser is one of the fastest existing multilingual parser worldwide (4000 8000 tokens/sec.). In order to ease model design for multilingual settings, we currently study efficient feature selection procedures for automating model adaptation to new languages.

The ongoing investigation aims to integrate continuous semantic representations into the model such as word embeddings in order to leverage data sparsity and estimation issues recurrent in lexicalized parsing. To this end we study neural-network-based architectures for structured phrase structure parsing.

6.5. Playing with DyALog-based parsers

Participant: Éric Villemonte de La Clergerie.

Éric de la Clergerie has continued the development of DyALog-SR, a transition-based dependency parser running on top of DyALog and initiated in 2013 to participate to SPMRL'2013. Thanks to DyALog's tabulation functionalities, this parser implements a dynamic programming algorithm to explore larger search space through the use of beams.

In order to participate to SemEval'14 Task on "broad coverage semantic dependency parsing", DyALog-SR was extended to handle non-connected dependency graphs rather than standard dependency trees. This was achieved by considering a richer set of transitions, besides the usual Shift and Reduce transitions. However, while working, this extended set of transitions was not ensuring the expected gains when using beams. The issue was finally solved after long investigations, with the identification of multiple causes. One of them was related to the fact that transition paths of various lengths may lead to a final state. In consequence, a noop transition was added to compensate on shorter paths.

A second axe of work was a thorough use of DyALog-SR over the French TreeBank (FTB) to compare its performances to those published for other parsers. By enriching its set of features and improving the update strategy of the perceptron-based statistical model of DyALog-SR, we were able to reach state-of-the-art results.

However, the best results were obtained by coupling DyALog-SR with FRMG, our large-coverage French grammar (derived from a meta-grammar). The results from FRMG were used as features to guide the statistical DyALog-SR parser. This innovative step proved to provide us with the best results published so far for the FTB (over 90% of Labeled Attachement Score [LAS] over the test part of the FTB) [41].

The improvements of FRMG was pursued in 2014, at the level of the underlying meta-grammar (to extend its coverage over 96% on the FTB) but also by adapting the statistical models developed for DYALOG-SR (in replacement of older and slower SQLite-based models).

6.6. Multiword expressions and statistical parsing

Participants: Sarah Beniamine, Marie-Hélène Candito, Benoît Sagot, Djamé Seddah.

Multi-word expressions recognition (MWE recognition) and syntactic parsing are two tasks that have been extensively investigated. Yet, systems combining both tasks have been rather rare. In particulat, works on parsing have tended to use training and test data with gold MWEs (generally with each MWE) merged into one token. In 2013, Djamé Seddah led the organization of the first shared task on statistical parsing Morphologically Rich Languages (SPMRL) [127], hosted by the fourth SPMRL workshop. The primary goal of this shared task was to bring forward work on parsing morphologically ambiguous input in both dependency and constituency parsing, and to show the state of the art for MRLs. The shared task proposed a data set for 9 languages. The French part of this data set is particular, in that it uses a representation combining MWEs and syntax, which allows to investigate techniques for performing parsing and MWE recognition. A first system was proposed for the dependency parsing track of the Shared Task, in collaboration with Matthieu Constant (LIGM, Université Marne-la-Vallée) [74]. This work investigates pipeline and joint architecture for both tasks. In 2014, Marie Candito and Matthieu Constant continued that line of work [2], focusing on using an alternative representation of syntactically regular MWEs, which captures their syntactic internal structure. The objective of such representation was two fold. First, it is well-known that the MWE status is not clear-cut, and that MWE status can hold due to syntactic and/or semantic criteria. In particular, syntactically regular MWEs exhibit various degrees of semantic non-compositionality. For such MWEs, an atomic representation fails to capture internal partial semantic composition, and also fails to take advantage of the internal syntactic regularity. Indeed, one hypothesis of this work was that augmenting the regularity of the syntactic representations could help parsing. The results of this work is that while this hypothesis could not be verified, the resulting system has comparable performance to that of previous works on this dataset, but it has the advantage of predicting both syntactic dependencies and the internal structure of MWEs, a crucial feature to capture the various degrees of semantic compositionality of MWEs.

In the same time, Sarah Beniamine and Benoît Sagot also investigated the use of internal regular structures for MWEs, yet for *syntagmatic* syntactic parsing. The objective is to guide a parser with predicted MWEs, while keeping a regular syntactic representation.

6.7. Graph-based approaches for deep-syntatic and semantic parsing

Participants: Corentin Ribeyre, Djamé Seddah, Éric Villemonte de La Clergerie.

With most state-of-the-art statistical parsers routinely crossing a ninety percent performance plateau in capturing tree structures, the question of *what next* crucially arises. Most of the structures used to train current parsing models are degraded versions of a more informative data set: the Wall Street journal section of the Penn treebank ([91]) which is often stripped from its richer set of annotations (i.e. traces and functional labels are removed), while, for reasons of efficiency and availability, projective dependency trees are often given preference over richer graph structures [96], [107]. This led to the emergence of *surface* syntax-based parsers [70], [97], [100] whose output cannot by itself be used to extract full-fledged predicate argument-structures. For example, control verb constructions, it-cleft structures, argument sharing in ellipsis coordination, etc. are among the phenomena requiring a graph to be properly accounted for. The dichotomy between what can usually be parsed with high accuracy and what lies in the deeper syntactic description has initiated a line of research devoted to closing the gap between surface syntax and richer structures.

At Alpage, we built our work on the widely known transition-based parsing approach [95], which is state-of-the-art to parse surfacic syntatic trees [141]. Shift-reduce transition-based parsers essentially rely on *configurations* formed of a stack and a buffer, with stack transitions used to move from a configuration to the next one, until reaching a final configuration.

6.8. English Broad-coverage Semantic Dependency Parsing

Participants: Corentin Ribeyre, Djamé Seddah, Éric Villemonte de La Clergerie.

We successfully tested our graph-based approach described in Section 6.7 on a shared task on broad-coverage semantic dependency parsing part of the International Workshop on Semantic Evaluation (SemEval 2014, [99]). We were given three resources, which constitute parallel semantic annotations over the same common text (the Penn Treebank (PTB), [91]). The first one is part of the tectogrammatical layer of the Prague Czech-English Dependency Treebank, the second one is the reduction of the Minimal Recursion Semantics, available through the HPSG annotation of the PTB, into bi-lexical dependencies [82]. Finally, the third one is the predicate-argument structures extracted from the Enju Parser [131]. The shared task consisted of two tracks: a closed one where we needed to use these three resources only and an open one, where we could use whatever we needed to produce the best semantic representations.

At Alpage, we developed two semantic parsers: The first one is based on a previous work on DAG parsing [107] and the second one on the FRMG surfacic syntactic parser [133]. We use two parsers to assess the validity of our approach. The top performing models we submitted used a mix of syntactic features (tree fragments from a constituent syntactic parser [100], dependencies from a syntactic parser [58], elementary spinal trees using a spine grammar [126], etc.) to improve our results. Our intuition is that syntax and semantic are not independent of each other and using syntax could improve semantic parsing. Our systems performs well and were able to compete with the top performers. Those systems, as well as the software needed to parse these new data sets, are already available.

6.9. Development of syntactic and deep-syntactic treebanks: Extending our Coverage

Participants: Djamé Seddah, Marie-Hélène Candito, Corentin Ribeyre, Benoît Sagot, Éric Villemonte de La Clergerie.

Taking its roots in the teams that initiated the first syntactically annotated the French Treebank, the first metagrammar compiler and one of the best wide coverage grammars, Alpage has a strong tendency to focus on creating pioneer resources that serve both to extend our linguistics knowledge and to nurture accurate parsing models. Recently, we focused on extending the lexical coverage of our parsers using semi-supervized techniques (see above) built on edited texts. In order to evaluate these models, we built the first free out-domain treebank for French (the Sequoia treebank, [69]) covering various domains such as Wikipedia, Europarl and bio medical texts on which we established the state-of-the-art. Exploring other kind of texts (speech, user generated content), we faced however various issues inherently tied to the nature of these productions. Syntactic divergences from the norm are actually prominent and are a severe bottleneck for any data driven parsing model. Simply because a structure not present in a training set cannot be reproduced. This analysis naturally occurred as a side effect of our experiments in parsing social media texts. Actually, the first version of the French Social Media Bank (FSMB) was conceived as a stress test for our tool chains (tokenization, tagging, parsing). Our recent experiments showed that to reach a decent performance plateau, we need to include some of the target data into our training set. Focusing on processing direct questions and social media texts, we built two treebanks of about 2,500 sentences each: one devoted to questions and one built to extend the FSMB [0]. These initatives are funded by the Labex EFL.

- The French Social Media Bank 2.0: We are about to release the second part of the FSMB, 2600 sentences from Twitter, Facebook and other sources, with an extended annotation scheme able to describe more precisely the various phenomena at stakes in the social media text streams. To do so we extended our pre-processing chain (included and available in the MeLT tagger) to include a much more robust normalizer and tokenizer than the one we used to build the first version of the FSMB. The building phase being over, publications on this topics are on preparation.

[0] Let us note that the ever evolving nature of user generated content makes this a necessity.

- The French Question Bank: The building of a treebank made solely of questions comes from the simple fact that in both the FTB and the Sequoia treebank, there's only 150 direct questions. Making the parsing of such constructions extremely difficult for our data driven parsers. Following our now classical methodology, we selected more than 3200 sentences coming from governmental sources, from the TREC ressources – allowing to have a strong set of aligned sentences with the English ressources – and from social media sources as well. In the case of the TREC part, those are the questions used by [85], which allows some potentially interesting cross-language experiments. Unlike in the English Question Bank, phrasal-movement are annotated with functional paths and not traces. This allows to maintain a strong compatibility with the FTB annotation scheme. Our Question bank is the only resources of its kind for any other languages than English.

Both ressources are available in constituency and dependency. The later being still verified for the FSMB 2.0.

Note that we just started another annotation campaign aiming at adding a deep syntax layer to these two data sets, following the Deep Sequoia as presented above. These resources will prove invaluable to building a robust data driven syntax to semantic interface.

In the same time, Alpage collaborated with the Nancy-based Inria team Sémagramme in the domain of deep syntax analysis. Deep Syntax is intended as an intermediary level of representation, at the interface between syntax and semantics, which partly abtracts away from syntactic variation, and aims at providing the canonical grammatical functions of predicates. This means for instance neutralizing diathesis alternation and making explicit argument sharing, such as occurring for infinitival verbs. The advantage of a deep syntactic representation is to provide a more regular representation to serve as basis for semantic analysis. Note though it is computationally more complex, as we switch from surface syntactic trees to deep syntactic graphs, since shared arguments are made explicit.

We collaboratively defined a deep syntactic representation scheme for French and built a gold deep syntactic treebank [21], [43]. More precisely, each team used an automatic surface-to-deep syntax converter module, applied it on the Sequoia corpus (already annotated for surface syntax), and manually corrected it. Remaining differences were collaboratively adjudicated. The surface-to-deep syntax converter tool used by Alpage is built around the OGRE Graph Rewriting Engine built by Corentin Ribeyre [105].

The Deep Sequoia Treebank is too small to train a deep syntactic analyzer directly. In order to obtain more annotated data, we further used the surface-to-deep syntax converter to obtain predicted (non validated) deep syntactic representations for the French Treebank [36], which is much bigger than the Sequoia treebank (more than 18.000 sentences compared to $3,000$ sentences). We performed an evaluation of a small subset of the resulting deep syntactic graphs. The high level of performance we obtained (more than 98% of F-score in labeled dependencies recovery task) which suggests that the deep syntax version of the French Treebank can be used as pseudo-gold data to train deep syntactic parsers, or to extract syntactic lexicons augmented with quantitative information.

6.10. Towards a French FrameNet

Participants: Marie-Hélène Candito, Marianne Djemaa, Benoît Sagot.

The ASFALDA project [0] is an ANR project coordinated by Marie Candito. 5 partners collaborate on the project, on top of Alpage : the Laboratoire d'Informatique Fondamentale de Marseille(LIF), the Laboratoire de Linguistique Formelle (LLF), the MELODI team (IRIT - Toulouse) and the CEA-List. It is a three-year project which started in October 2012, with the objective of building semantic resources (generalizations over predicates and over the semantic arguments of predicates) and a corresponding semantic analyzer for French. We chose to build on the work resulting from the FrameNet project [57], [0] which provides a structured set of prototypical situations, called *frames*, along with a semantic characterization of the participants of these situations (called *frame elements*). The resulting resources will consist of :

[0] https://sites.google.com/site/anrasfalda/
[0] https://framenet.icsi.berkeley.edu/

1. a French lexicon in which lexical units are associated to FrameNet frames,
2. a semantic annotation layer added on top of existing syntactic French treebanks
3. and a frame-based semantic analyzer, focused on joint models for syntactic and semantic analysis.

In 2014, we first finished the work on the lexicon, which was started in 2013 [19]. The step 2 (semantic annotations on top of syntactic representations) is ongoing :

- We wrote the annotation guide. In particular Marianne Djemaa focused on how to annotate phenomenon known to exhibit syntax/semantic divergences [42].
- We designed the annotation workflow and built an automatic pre-annotator, which proposes candidate semantic annotations that must be disambiguated manually.
- We started in july 2014 to manage six annotators, who were hired to perform the manual annotation phase.

6.11. Towards a morpho-semantic resource for French designed for Word Sense Disambiguation

Participant: Lucie Barque.

The most promising WSD methods are those relying on external knowledge resources [93] but semantic resources for French are scarse. Moreover, existing resources offer fine grained sense dictinctions that do not fit to WSD. Our aim is to provide the NLP community with a broad-coverage morpho-semantic lexicon for French that relies on coarse-grained sense distinctions for polysemic units. Preliminary results concern nouns, on which we have first focused because their semantic description, compared to verbs, crucially lacks (for information retrieval, for instance) and because the regular polysemy phenomenon (recurring cases of polysemy within semantic classes) mainly occurs in nominal semantic classes:

- We proposed a linguistically motivated description of general semantic labels for nouns, that will allow for coarse-grained sense distinctions [40]
- Regular polysemy of nouns that can denote an event or a participant of this event has also been described for a large number of French nouns in [12]

6.12. Development of Verb∋net

Participants: Laurence Danlos, Quentin Pradet.

VerbNet is an English lexical resources for verbs, which is internationally known and widely used in numerous NLP applications [89]. Verb∋net is a French adaptation of this resource. It is semi-automatically developed thanks to the use of two French existing resources created in the 70's: LG, Lexique-Grammaire developed at LADL under the supervision of Maurice Gross, and LVF, Lexique des verbs du français by Dubois and Dubois-Charlier. The idea is to map English classes, which gather verbs with a common syntactic and semantic behavior, into classes of LG and LVF, then to manually adapt the syntactic frames according to French grammar while keeping the thematic roles and the semantic information, [35], [28]. This work is currently under progress in collaboration with Takuya Nakamura (Institut Gaspard Monge) and the resource should be freely available in 2015.

6.13. Development of FDTB1

Participants: Laurence Danlos, Margot Colinet, Jacques Steinlin.

FDTB1 is the first step towards the creation of the French Discourse Tree Bank (FDTB) with a discourse layer on top of the syntactic one which is available in the French Tree Bank (FTB). In this first step, we have identified all the words or phrases in the corpus that are used as “discourse connectives”. The methodology was the following: first, we highlighted all the items in the corpus that are recorded in LexConn [106], a lexicon of French connectives with 350 items, next we eliminated some of these items with the following criteria:

1. first, we filtered out the LexConn items that are annotated in FTB with parts of speech incompatible with a connective use, e.g. *bref* annotated as *Adj* instead of *Adv*, *en fait* annotated as *Pro V* instead of (compound) *Adv*;
2. second, as we lay down for theoretical and pratical reasons that elementary arguments of connectives must be clauses or VPs, we filtered out e.g. LexConn prepositions that introduce NPs;
3. last, we filtered out LexConn prepositions and adverbials with a non-discursive function.

The last criterion requires a manual work contrarily to the two others. For example the preposition *pour (to)*, is ambiguous between a connective use *(Fred s'est dépeché pour être à la gare à 17h (Fred hurried to be at the station at 17h))* and a preposition introducing a complement *(Fred s'est dépeché pour aller à la gare (Fred hurried to go to the station))*, and the disambiguation between the two uses is subtle and so the topic of a long paper [22], whose results have been used to enhance Le*fff*, [44].

The FDTB corpus contains 18 535 sentences and FDTB1 identifies 9 833 discourse connectives. This ressource is freely available.

6.14. Discourse Parsing

Participants: Laurence Danlos, Chloé Braud.

Discourse parsing goal is to reflect the rhetorical structure of a document, how pieces of text are linked in order to form a coherent document. Understanding such links could benefits to several other natural language applications (summarization, language generation, information extraction...). A discourse parser corresponds to two major subtasks: a segmentation step wherein discourse units (DUs) are extracted, and a parsing step wherein these DUs are (recursively) related through “discourse (rhetorical) relations”. The more difficult task in discourse parsing is the labeling of the relations between DUs, especially when no so-called connective overtly marks the relation (we then talk about implicit relations as opposed to explicit ones). In her PhD work, Chloé Braud develops a discourse relation classifier, carrying experiments on French and English. Focusing on the problem on implicit relation identification, this work tries to tackle the lack of manually annotated data, a discourse specific difficulty, by exploiting the similarities between explicit and implicit relations. In 2014, this work lead to systems based on domain adaptation methods [18], [13], demonstrating improvements on the French corpus Annodis [56].

6.15. Multilingual and cross-lingual terminology extraction

Participants: Valérie Hanoka, Benoît Sagot.

Language diversity spans more than 7000 languages. Among them, 24 macrolanguages [0] have at least 50 million first-language speakers. Traditional terminology techniques, which are mostly based on language-dependent linguistic tools (part of speech tagging, phrase chunking) requires a considerable effort to be developed for a new language. This effort is likely to be even more critical if the term extraction is to be based on noisy text (i.e. displaying linguistic creativity, spelling errors and ungrammatical sentences). In this context, the need has arisen to examine the issue of a less language-specific method for term extraction.

To that end, our approach take advantage of existing language typologies in order to alleviate for the lack of language-dependent linguistic processing. We based our reflexions and experiments on a sample of 7 typologically different language: Arabic, Chinese, English, French, German, Polish and Turkish.

[0] A macrolanguage is defined as "multiple, closely related individual languages that are deemed in some usage contexts to be a single language" in the ISO 639-3 standard.

As a starting point, we considered the minimal textual preprocessing (character normalization, segmentation) needed to allow for a comprehensive multilingual approach to automatic term extraction. In order to gain further insight on the influence of the morphology for term extraction, we examined the impact of the deletion of selected morphological information on words of morphologically rich languages.
For the different settings, models based on Conditional Random Fields (CRF) have been trained on existing gold data. We proposed an adapted version of the evaluation algorithm of [94] able to issue terminological scores for all the language of our sample. The scores thus obtained allowed to identify the best experimental setting for each language tested.

The results were surprising in two ways: First, the cross-lingual [0] application of models works well (the best cross-lingual models' accuracies range from 0.8% to 0.97%). Secondly, the languages which makes the overall best cross-lingual models are those who have the richest morphology (i.e: Turkish).

Finally, we developped and used a multilingual translation graph [32] to extend the multilingual terminology obtained using two methods: those presented in [83] and a more formal one, based on a simulated annealing clustering algorithm.

6.16. Word order variation in Old French

Participants: Benoit Crabbé, Alexandra Simonenko, Benoît Sagot.

As participant of the strand *Experimental Grammar* of the Labex EFL project *Empirical Foundations of Linguistics*[0] we study word order issues on Old French and more specifically the relative ordering of complements of ditransitive verbs. The inquiry seeks to identify several factors influencing the ordering of Old French complementation in different texts (varying in dates and genres) by carrying quantitative and statistical work from annotated Old French data. [0]

The quantitative results will be compared with what is known from corpus studies on the relative ordering of subject and complement in Old French [90]. It will also be compared to the quantitative results obtained on the relative ordering of complements of ditransitive verbs in Modern French [8] and modern English [64]. This comparative perspective is expected to provide new insights on French language evolution.

6.17. Cross linguistic factors governing word order

Participants: Benoit Crabbé, Kristina Gulordava.

In many languages, flexible word order often has a pragmatic role and marks the introduction of new information, a focus or a topic shift. Other cases of language- internal word order variation are alternations between two options such as *Mary gave John a book* and *Mary gave a book to John*, which are conditioned on syntactic and semantic factors such as the complexity of the constituents (as in *Mary gave John a book she had read ten times*), their animacy or the meaning of the verb [63].

One of the goals of this module is to investigate the connection between the quantitative aspects of word order variation across languages and the quantitative aspects of word order variation within a language. We study the corresponding patterns in language-internal variation by looking at the syntactically annotated corpora of various languages. Focusing on the variation of the internal word order of the noun-phrase as a case study, we explore to which extent a computational corpus-based analysis can provide new evidence not only for empirical, but also for theoretical linguistic research.

6.18. Anaphoricity detection and coreference resolution

Participant: Emmanuel Lassalle.

[0] A model trained on data of one language and applied to data of another language
[0] www.labex-efl.org
[0] SRCMF corpus: http://srcmf.org/; MCVF: http://www.voies.uottawa.ca

Resolving coreference in a text, that is, partitioning mentions (noun phrases, verbs, etc) into referential entities, is a challenging task in NLP leading to many different approaches. Anaphoricity detection, on the other hand, consists in deciding whether a mention is anaphoric (aka discourse-old) or non-anaphoric (discourse-new). This task is strongly related to coreference resolution and has been mainly addressed as a preliminary task to solve, leading to pipeline architectures.

A first line of work compares several methods for learning latent structures encoding coreference clusters that optionally take into account very accurate constraints on mention pairs. We study the relationship between standard decoding strategies used with pairwise models and those used with structured learning of latent structures, providing both topological and empirical comparisons. We also show that further gains can be obtained by the addition of pairwise constraints. Our experiments on the CoNLL-2012 dataset show that our best system obtains state-of-the-art results, and significant gains compared to standard locally-trained models.

Our second line of work introduces a new structured model for learning anaphoricity detection and coreference resolution in a jointly. Specifically, we use a latent tree to represent the full coreference and anaphoric structure of a document at a global level, and we jointly learn the parameters of the two models using a version of the structured perceptron algorithm. This model is refined by the use of pairwise constraints, and our experiments on the CoNLL-2012 English datasets show large improvements in both coreference resolution and anaphoricity detection, compared to various competing architectures. Our best coreference system obtains a CoNLL score of 81.97 on gold mentions, which is to date the best score reported on this setting.

This work has been achieved in collaboration with Pascal Denis, a former Alpage member, now at Inria Lille-Nord-Europe (EPI Magnet).

ALPINES Project-Team

6. New Results

6.1. Highlights of the Year

We have released a version of FreeFem++ (v 3.33) which introduces new and important features related to high performance computing:

- Interface with PETSc library
- Interface with HPDDM (see above)
- improved interface with the parallel direct solver MUMPS

This release enables, for the first time, end-users to run the very same code on computers ranging from laptops to clusters and even large scale computers with thousands of computing nodes

6.2. Communication avoiding algorithms for dense linear algebra

Our group continues to work on algorithms for dense linear algebra operations that minimize communication. During this year we focused on improving the performance of communication avoiding QR factorization as well as designing algorithms that reduce communication on multilevel hierarhical platforms.

In [17] we focus on the QR factorization. The Tall-Skinny QR (TSQR) algorithm is more communication efficient than the standard Householder algorithm for QR decomposition of matrices with many more rows than columns. However, TSQR produces a different representation of the orthogonal factor and therefore requires more software development to support the new representation. Further, implicitly applying the orthogonal factor to the trailing matrix in the context of factoring a square matrix is more complicated and costly than with the Householder representation. We show how to perform TSQR and then reconstruct the Householder vector representation with the same asymptotic communication efficiency and little extra computational cost. We demonstrate the high performance and numerical stability of this algorithm both theoretically and empirically. The new Householder reconstruction algorithm allows us to design more efficient parallel QR algorithms, with significantly lower latency cost compared to Householder QR and lower bandwidth and latency costs compared with Communication-Avoiding QR (CAQR) algorithm. As a result, our final parallel QR algorithm outperforms ScaLAPACK and Elemental implementations of Householder QR and our implementation of CAQR on the Hopper Cray XE6 NERSC system.

In [18] we focus on performance predictions of multilevel communication optimal LU and QR factorizations on hierarchical platforms. This study focuses on the performance of two classical dense linear algebra algorithms, the LU and the QR factorizations, on multilevel hierarchical platforms. We first introduce a new model called Hierarchical Cluster Platform (HCP), encapsulating the characteristics of such platforms. The focus is set on reducing the communication requirements of studied algorithms at each level of the hierarchy. Lower bounds on communications are therefore extended with respect to the HCP model. We then introduce multilevel LU and QR algorithms tailored for those platforms, and provide a detailed performance analysis. We also provide a set of numerical experiments and performance predictions demonstrating the need for such algorithms on large platforms.

6.3. Enlarged Krylov methods

Krylov subspace methods are among the most practical and popular iterative methods today. They are polynomial iterative methods that aim to solve systems of linear equations ($Ax = b$) by finding a sequence of vectors $x_1, x_2, x_3, x_4, ..., x_k$ that minimizes some measure of error over the corresponding spaces $x_0 + \mathcal{K}_i(A, r_0), \quad i = 1, ..., k$ where $\mathcal{K}_i(A, r_0) = span\{r_0, Ar_0, A^2r_0, ..., A^{i-1}r_0\}$ is the Krylov subspace of dimension i, x_0 is the initial iterate, and r_0 is the initial residual. These methods are governed by Blas1 and

Blas2 operations as dot products and sparse matrix vector multiplications. Parallelizing dot products is constrained by communication since the performed computation is negligible. If the dot products are performed by one processor, then there is a need for a communication before and after the computation. In both cases, communication is a bottleneck. In [21] we introduce a new approach for reducing communication in Krylov subspace methods that consists of enlarging the Krylov subspace by a maximum of t vectors per iteration, based on the domain decomposition of the graph of A. The obtained enlarged Krylov subspace $\mathcal{K}_{t,k}(A, r_0)$ is a superset of the Krylov subspace $\mathcal{K}_k(A, r_0)$, $\mathcal{K}_k(A, r_0) \subset \mathcal{K}_{t,k+1}(A, r_0)$. Thus it is possible to search for the solution of the system $Ax = b$ in $\mathcal{K}_{t,k}(A, r_0)$ instead of $\mathcal{K}_k(A, r_0)$. Moreover, we show that the enlarged Krylov projection subspace methods lead to faster convergence in terms of iterations and parallelizable algorithms with less communication, with respect to Krylov methods.

6.4. Algebraic preconditioners

Our work focused on the design of robust algebraic preconditioners and domain decomposition methods to accelerate the convergence of iterative methods.

In [8] we introduce the block filtering decomposition, a new preconditioning technique that is suitable for matrices arising from the discretization of a system of PDEs on unstructured grids. The preconditioner satisfies a so-called filtering property, which ensures that the input matrix is identical with the preconditioner on a given filtering vector. This vector is chosen to alleviate the effect of low frequency modes on convergence and so decrease or eliminate the plateau which is often observed in the convergence of iterative methods. In particular, the paper presents a general approach that allows to ensure that the filtering condition is satisfied in a matrix decomposition. The input matrix can have an arbitrary sparse structure. Hence, it can be reordered using nested dissection, to allow a parallel computation of the preconditioner and of the iterative process. We present experimental results that demonstrate the efficiency of the proposed preconditioner on a set of matrices arising from the discretization of partial differential equations on two-dimensional and three-dimensional grids. We also show that the numerical efficiency of the preconditioner does not suffer from the reordering of the unknowns for the matrices in our test set, which can have highly heterogeneous and anisotropic coefficients.

In [9] we discuss the usage of overlapping techniques for improving the convergence of preconditioners based on incomplete factorizations. To enable parallelism, these preconditioners are usually applied after the input matrix is permuted into nested bordered block diagonal form. We use k-way partitioning with vertex separator (KPVS) to recursively partitions the corresponding graph of the input matrix into k subgraphs using a subset of its vertices called separators. In the case where $k = 2$, it is called nested dissection. The overlapping technique is then based on algebraically extending the associated subdomains of these subgraphs and their corresponding separators obtained from KPVS by their direct neighbours. This approach is known to accelerate the convergence of domain decomposition methods, where the input matrix is partitioned into a number of independent subdomains using k-way graph partitioning, a different graph decomposition technique. We discuss the effect of the overlapping technique on the convergence of two classes of preconditioners, based on nested factorization and block incomplete LDU factorization.

In [22] we introduce LORASC, a robust algebraic preconditioner for solving sparse linear systems of equations involving symmetric and positive definite matrices. The graph of the input matrix is partitioned by using k-way partitioning with vertex separators into N disjoint domains and a separator formed by the vertices connecting the N domains. The obtained permuted matrix has a block arrow structure. The preconditioner relies on the Cholesky factorization of the first N diagonal blocks and on approximating the Schur complement corresponding to the separator block. The approximation of the Schur complement involves the factorization of the last diagonal block and a low rank correction obtained by solving a generalized eigenvalue problem or a randomized algorithm. The preconditioner can be build and applied in parallel. Numerical results on a set of matrices arising from the discretization by the finite element method of linear elasticity models illustrate the robusteness and the efficiency of our preconditioner.

The Helmholtz equation governing wave propagation and scattering phenomena is difficult to solve numerically. Its discretization with piecewise linear finite elements results in typically large linear systems of equations. The inherently parallel domain decomposition methods constitute hence a promising class of preconditioners. An essential element of these methods is a good coarse space. Here, the Helmholtz equation presents a particular challenge, as even slight deviations from the optimal choice can be devastating.

In [5], we present a coarse space that is based on local eigenproblems involving the Dirichlet-to-Neumann operator. Our construction is completely automatic, ensuring good convergence rates without the need for parameter tuning. Moreover, it naturally respects local variations in the wave number and is hence suited also for heterogeneous Helmholtz problems. The resulting method is parallel by design and its efficiency is demonstrated on 2D homogeneous and heterogeneous numerical examples.

Coarse spaces are instrumental in obtaining scalability for domain decomposition methods for partial differential equations (PDEs). However, it is known that most popular choices of coarse spaces perform rather weakly in the presence of heterogeneities in the PDE coefficients, especially for systems of PDEs. In [12], we introduce in a variational setting a new coarse space that is robust even when there are such heterogeneities. We achieve this by solving local generalized eigenvalue problems in the overlaps of subdomains that isolate the terms responsible for slow convergence. We prove a general theoretical result that rigorously establishes the robustness of the new coarse space and give some numerical examples on two and three dimensional heterogeneous PDEs and systems of PDEs that confirm this property.

Multiphase, compositional porous media flow models lead to the solution of highly heterogeneous systems of Partial Differential Equations (PDEs). In [7] , we focus on overlapping Schwarz type methods on parallel computers and on multiscale methods. We recall a coarse space that is robust even when there are such heterogeneities. The two-level domain decomposition approach is compared to multiscale methods.

In [16], we investigate two-level preconditioners on the extended linear system arising from the domain decomposition method. The additive Schwarz method is used as a smoother, and the coarse grid space is constructed by using the Ritz vectors obtained in the Arnoldi process. The coarse grid space can be improved adaptively as the Ritz vectors become a better approximation of the eigenvectors. Numerical tests on the model problem demonstrate the efficiency.

6.5. New results related to FreeFem++

In [10], we propose an efficient algorithm for the numerical approximation of metrics, used for anisotropic mesh adaptation on triangular meshes with finite element computations. We derive the metrics from interpolation error estimates expressed in terms of higher order derivatives, for the $P-k$-Lagrange finite element, $k>1$. Numerical examples of mesh adaptation done using metrics computed with our Algorithm, and derived from higher order derivatives as error estimates, show that we obtain the right directions of anisotropy.

In [2], we consider a system of two reaction-dispersion equations with non constant parameters. Both equations are coupled through the boundary conditions. We propose a mixed variational formulation that leads to a non symmetric saddle-point problem. We prove its well-posedness. Then, we develop a stabilized mixed finite element discretization of this problem and establish optimal a priori error estimates.

In [15], we consider a model of soil water and nutrient transport with plant root uptake. The geometry of the plant root system is explicitly taken into account in the soil model. We first describe our modeling approach. Then, we introduce an adaptive mesh refinement procedure enabling us to accurately capture the geometry of the root system and small-scale phenomena in the rhizosphere. Finally, we present a domain decomposition technique for solving the problems arising from the soil model as well as some numerical results.

6.6. Auto adaptative algorithms

In [29], we develop an adaptive version of the inexact Uzawa algorithm applied to finite element discretizations of the linear Stokes problem. We base our developments on an equilibrated flux a posteriori error estimate distinguishing the different error components, namely the discretization error component, the inner algebraic solver error component, and the outer Uzawa iteration error component. On each outer Uzawa and inner linear algebraic solver iteration, we prove that our estimate gives a guaranteed upper bound on the total error, as well as a polynomial-degree-robust local efficiency. Our adaptive inexact algorithm stops the outer Uzawa iteration and the inner linear algebraic solver iteration when the Uzawa error component, respectively the algebraic solver error component, do not have a significant influence on the total error. The developed framework covers all standard conforming and conforming stabilized finite element methods. The implementation into the FreeFem++ programming language is invoked and two numerical examples showcase the performance of our adaptive strategy.

6.7. Spectrum for a small inclusion of negative material

We studied a spectral problem $(\mathcal{P}^\delta)$ for a diffusion like equation in a 3D domain Ω. The main originality here lies in the presence of a parameter σ^δ, whose sign changes on Ω, in the principal part of the operator we consider. More precisely, σ^δ is positive on Ω except in a small inclusion of size $\delta > 0$. Because of the sign-change of σ^δ, for all $\delta > 0$ the spectrum of $(\mathcal{P}^\delta)$ consists of two sequences converging to $+\infty$ and $-\infty$. However, at the limit $\delta = 0$, the small inclusion vanishes so that there should only remain positive spectrum for $(\mathcal{P}^\delta)$. What happens to the negative spectrum? In this paper, we prove that the positive spectrum of $(\mathcal{P}^\delta)$ tends to the spectrum of the problem without the small inclusion. On the other hand, we establish that each negative eigenvalue of $(\mathcal{P}^\delta)$ behaves like $\delta^{-2}\mu$ for some constant $\mu < 0$. We also show that the eigenvectors associated with the negative eigenvalues are localized around the small inclusion. We end the article providing 2D numerical experiments illustrating these results.

6.8. Stability of electromagnetic cavities perturbed by small perfectly conducting inclusions

We consider an electromagnetic wave propagation problem in harmonic regime in a bounded cavity, in the case where the medium of propagation contains small perfectly conducting inclusions. We prove that the solution to this problem depends continuously on the data in a uniform manner with respect to the size of the inclusions.

6.9. Integral equations for acoustic scattering by partially impenetrable composite objects

We study direct first-kind boundary integral equations arising from transmission problems for the Helmholtz equation with piecewise constant coefficients and Dirichlet boundary conditions imposed on a closed surface. We identify necessary and sufficient conditions for the occurrence of so-called spurious resonances, that is, the failure of the boundary integral equations to possess unique solutions.

Following [A. Buffa and R. Hiptmair, Numer Math, 100, 1–19 (2005)] we propose a modified version of the boundary integral equations that is immune to spurious resonances. Via a gap construction it will serve as the basis for a universally well-posed stabilized global multi-trace formulation that generalizes the method of [X. Claeys and R. Hiptmair, Commun Pure and Appl Math, 66, 1163–1201 (2013)] to situations with Dirichlet boundary conditions.

6.10. Application domain: data analysis in astrophysics

One of the application domain on which our algorithms are validated is data analysis in astrophysics. Estimation of the sky signal from sequences of time order data is one of the key steps in the Cosmic Microwave Background (CMB) data analysis, commonly referred to as the map-making problem. Some of the most popular and general methods proposed for this problem involve solving generalised least squares (GLS) equations with non-diagonal noise weights given by a block-diagonal matrix with Toeplitz blocks. In [14] we study new map-making solvers potentially suitable for applications to the largest, anticipated data sets. They are based on iterative conjugate gradient (CG) approaches enhanced with novel, parallel, two-level preconditioners (2lvl-PCG). We apply the proposed solvers to examples of simulated, non-polarised and polarised CMB observations and a set of idealised scanning strategies with a sky coverage ranging from nearly a full sky down to small sky patches. We discuss in detail their implementation for massively parallel computational platforms and their performance for a broad range of parameters characterising the simulated data sets. We find that our best new solver can outperform carefully optimised, standard solvers as used today, by as much as a factor of 5 in terms of the convergence rate and a factor of 4 in terms of the time to solution, and does so without increasing significantly the memory consumption or the volume of inter-processor communication. The performance of the new algorithms is also found to be more stable, robust and less dependent on specific characteristics of the analysed data set.We therefore conclude that the proposed approaches are well suited to address successfully challenges posed by new and forthcoming CMB data sets.

Spherical Harmonic Transforms (SHT) are at the heart of many scientific and practical applications ranging from climate modelling to cosmological observations. In many of these areas new, cutting-edge science goals have been recently proposed requiring simulations and analyses of experimental or observational data at very high resolutions and of unprecedented volumes. Both these aspects pose formidable challenge for the currently existing implementations of the transforms.

In [13] we describe parallel algorithms for computing SHT with two variants of intra-node parallelism appropriate for novel supercomputer architectures, multi-core processors and Graphic Processing Units (GPU). It also discusses their performance, alone and embedded within a top-level, MPI-based parallelisation layer ported from the S^2HAT library, in terms of their accuracy, overall efficiency and scalability. We show that our inverse SHT run on GeForce 400 Series GPUs equipped with latest CUDA architecture ("Fermi") outperforms the state of the art implementation for a multi-core processor executed on a current Intel Core i7-2600K. Furthermore, we show that an MPI/CUDA version of the inverse transform run on a cluster of 128 Nvidia Tesla S1070 is as much as 3 times faster than the hybrid MPI/OpenMP version executed on the same number of quad-core processors Intel Nehalem for problem sizes motivated by our target applications. Performance of the direct transforms is however found to be at the best comparable in these cases. We discuss in detail the algorithmic solutions devised for the major steps involved in the transforms calculation, emphasising those with a major impact on their overall performance, and elucidates the sources of the dichotomy between the direct and the inverse operations.

AMIB Project-Team

5. New Results

5.1. RNA

To mitigate the current absence of a selective scientific event dedicated to RNA computational biology, impeding the dissemination of recent methodological results, AMIB members have participated in the creation of the *Computational Methods for Structural RNAs* workshops (CMSR'14). This first installment of the event was hosted in Strasbourg as a workshop of the 2014 edition of European Conference on Computational Biology. Its proceedings were published by McGill University [33], and extended versions of selected articles were invited to appear in the *Journal of Computational Biology*.

5.1.1. RNA visualization

The field of RNA visualization is now rich with multiple tools that accommodate different needs, arising from a variety of application contexts. In order to help end-users navigate through the jungle of available options, Y. Ponty and F. Leclerc (IGM, Univ. Paris-Sud) have contributed a review of existing tools, and illustrate their usage to address a collection of typical use-cases [35].

5.1.2. RNA design and structures

The past couple of years have seen the multiplication of heuristic or exponential time algorithms for the RNA design problem. This situation motivates a survey, which s currently lacking, that would focus on the relative merits of existing algorithms, and assess their applicability towards the typical goals of synthetic biology. Such an objective evaluation is at the core of the PhD project of Vincent Le Gallic, which was started in September 2014.

With Antoine Soulé, a PhD student of J-M Steyaert and J. Waldispühl (McGill), a comparative study of the various softwares for the inverse RNA folding problem is under revision and a new version of RNAMUTANT in the langage GAP-L with enrichment has been designed.

Besides, we have published a general survey on RNA structure comparison [9].

5.1.3. RNA splicing regulation

RNA splicing is a modification of the nascent pre-messenger RNA (pre-mRNA) transcript in which introns are removed and exons are joined. The U2AF heterodimer protein has been well studied for its role in defining functional 3' splice sites in pre-mRNA splicing, but multiple critical problems are still outstanding, including the functional impact of their cancer-associated mutations. In collaboration with Xiang-Dong Fu's groups in San Diego and Wuhan, , through genome-wide analysis of U2AF-RNA interactions, we reported in [16] that U2AF has the capacity to define 88% of functional 3' splice sites in the human genome. Numerous U2AF binding events also occur in other genomic locations, and metagene and minigene analysis suggests that upstream intronic binding events interfere with the immediate downstream 3' splice site associated with either the alternative exon to cause exon skipping or competing constitutive exon to induce inclusion of the alternative exon.

5.1.4. RNA 3D structure modelling

Conformational diversity for RNA ensemble analyses is often provided by sophisticated molecular dynamics simulations. Long trajectories with specialized force fields on dedicated supercomputers are required to adequately sample conformational space, limiting ensemble analyses to modestly-sized RNA molecules. To avoid these limitations, we developed an efficient conformational sampling procedure, Kino-geometric sampling for RNA (KGSrna), which can report on ensembles of RNA molecular conformations orders of magnitude faster than MD simulations. In the KGSrna model, the RNA molecule is represented with rotatable, single bonds as degrees-of-freedom and groups of atoms as rigid bodies. In this representation, non-covalent bonds form distance constraints, which create nested, closed cycles in a rooted spanning tree. Torsional degrees-of-freedom in a closed ring demand carefully coordinated changes to avoid breaking the non-covalent bond, which greatly reduces the conformational flexibility. The reduced flexibility from a network of nested, closed rings consequently deforms the biomolecule along preferred directions on the conformational landscape. This new procedures projects degrees-of-freedom onto a lower-dimensional subspace of the conformation space, in which the geometries of the non-covalent bonds are maintained exactly under conformational perturbation. The dimensionality reduction additionally enables efficient exploration of conformational space and reduces the risk of overfitting sparse experimental data. Kinogeometric sampling of 3D RNA models can recover the conformational landscape encoded by proton chemical shifts in solution and is thus of great help to interpret NMR experimental data [11]. The computational efficiency of this approach, combined to its inherent parallel nature could also be adapted to model large assemblies on parallel platforms.

Our expertise was also essential in modelling junction of the RNA structure of a large biomolecule of interest, the tmRNA so as to study its interaction with the SmpB protein. Results obtained in collaboration with experiementalists, mainly P. Vachette at IBBMC and S. Nonin-Lecomte at the LCRB were made available in [15].

5.2. Sequences

5.2.1. Random generation

In collaboration with the Simon Fraser University (Vancouver, Canada), we have explored a random generation strategy, under a Boltzmann distribution, to assess the robustness of predicted adjacencies in ancestral genomes using a parsimony-based approach. The sampling algorithm was used to estimate the Boltzmann probability of ancestral adjacencies, which was then used as a filter to weed out unsupported predictions, leading to the resolution of a large number of syntenic inconsistencies [23].

5.2.2. Combinatorics of motifs

An algorithm for pvalue computation has been proposed in [40] that takes into account a Hiddden Markov Model and an implementation, SUFPREF, has been realized (http://server2.lpm.org.ru/bio).

Combinatorics of clumps have been extensively studied, leading to the definition of the so-called *canonic clumps*. It is shown in [26] that they contain the necessary information needed to calculate, approximate, and study probabilities of occurrences and asymptotics. This motivates the development of a *clump automaton*. It allows for a derivation of pvalues, decreasing the space and time complexity of the generating function approach or previous weighted automata. An extension to degenerate patterns is currently realized and implemented in a collaboration with J. Holub (Praha U.) and E. Furletova (IMPB).

During her master thesis at King's College, A. Héliou and collaborators designed the first linear-time and linear-space algorithm for computing all minimal absent words based on the suffix array [6]. In a typical application, one would be interested in computing minimal absent words to compare and study genomes in linear time by considering this negative information.

In a collaboration with AlFarabi University, where M. Régnier acts as a foreign co-advisor), word statistics were used to identify mRNA targets for miRNAs involved in various cancers [7].

5.2.3. Prediction and functional annotation of ortholog groups of proteins

In comparative genomics, orthologs are used to transfer annotation from genes already characterized to newly sequenced genomes. Many methods have been developed for finding orthologs in sets of genomes. However, the application of different methods on the same proteome set can lead to distinct orthology predictions.

In [38], [14] we developed a method based on a meta-approach that is able to combine the results of several methods for orthologous group prediction. The purpose of this method is to produce better quality results by using the overlapping results obtained from several individual orthologous gene prediction procedures. Our method proceeds in two steps. The first aims to construct seeds for groups of orthologous genes; these seeds correspond to the exact overlaps between the results of all or several methods. In the second step, these seed groups are expanded by using HMM profiles.

We evaluated our method on two standard reference benchmarks, OrthoBench and Orthology Benchmark Service. Our method presents a higher level of accurately predicted groups than the individual input methods of orthologous group prediction. Moreover, our method increases the number of annotated orthologous pairs without decreasing the annotation quality compared to twelve state-of-the-art methods.

5.3. 3D Modelling and Interactions

5.3.1. Transmembrane proteins

Transmembrane beta-barrel proteins (TMB) account for 20 to 30% of identified proteins in a genome but, due to difficulties with standard experimental techniques, they are only 2% of the RCSB Protein Data Bank. As TMB perform many vital functions, the prediction of their structure is a challenge for life sciences, while the small number of known structures prohibits knowledge-based methods for structure prediction. We study and design algorithmic solutions addressing the secondary structure, an abstraction of the 3D conformation of a molecule, that only retains the contacts between its residues. As TMBs are strongly structured objects, model based methodologies [18] are an interesting alternative to conventional methods. The efficiently obtained 3D structures provide a good model for further 3D and interaction analyses.

5.3.2. 3D Interaction prediction

While protein-RNA complexes provide a wide range of essential functions in the cell, their atomic experimental structure solving is even more difficult than for proteins. Protein-RNA complexes provide a wide range of essential functions in the cell. Docking approaches that have been developed for proteins are often challenging to adapt for RNA because of its inherent flexibility and the structural data available being relatively scarce.We adapted the reference RosettaDock protocol for protein-RNA complexes both at the nucleotide and atomic levels. Using a genetic algorithm-based strategy, and a non-redundant protein-RNA dataset, we derived a RosettaDock scoring scheme able not only to discriminate but also score efficiently docking decoys. The approach proved to be both efficient and robust for generating and identifying suitable structures when applied to two protein-RNA docking benchmarks in both bound and unbound settings. It also compares well to existing strategies. This is the first approach that currently offers a multi-level optimized scoring approach integrated in a full docking suite, leading the way to adaptive fully flexible strategies [28], [12]. This work is part of the PhD thesis of Adrien Guilhot-Gaudeffroy. While the previously described approaches perform well in a rigid or semi-flexible docking setting, the generation of putative conformations for flexible molecules (sampling) is still a difficult question that has to be addressed in a multi-scale setting involving new algorithms. Docking these sampled conformations will also certainly require improvement in clustering approaches.

5.4. Data Integration

With the increasing popularity of scientific workflows, public and private repositories are gaining importance as a means to share, find, and reuse such workflows.

As the sizes of these repositories grow, methods to compare the scientific workflows stored in them become a necessity, for instance, to allow duplicate detection or similarity search. Scientific workflows are complex objects, and their comparison entails a number of distinct steps from comparing atomic elements to comparison of the workflows as a whole. Various studies have implemented methods for scientific workflow comparison and came up with often contradicting conclusions upon which algorithms work best. Comparing these results is cumbersome, as the original studies mixed different approaches for different steps and used different evaluation data and metrics.

In collaboration with members of the University of Humboldt (Berlin), we first contribute to the field [17] by (i) comparing in isolation different approaches taken at each step of scientific workflow comparison, reporting on an number of unexpected findings, (ii) investigating how these can best be combined into aggregated measures, and (iii) making available a gold standard of over 2000 similarity ratings contributed by 15 workflow experts on a corpus of 1500 workflows and re-implementations of all methods we evaluated. In this context, we have designed new approaches based on consensus ranking [21] to provide a consensus of the experts' answers.

Then, with members of the University of Pennsylvania, we have presented a novel and intuitive workflow similarity measure that is based on layer decomposition [27] (designed during the month SCB spent at UPenn). Layer decomposition accounts for the directed dataflow underlying scientific workflows, a property which has not been adequately considered in previous methods. We comparatively evaluate our algorithm using our gold standard and show that it a) delivers the best results for similarity search, b) has a much lower runtime than other, often highly complex competitors in structure-aware workflow comparison, and c) can be stacked easily with even faster, structure-agnostic approaches to further reduce runtime while retaining result quality.

Another way to make scientific workflows easier to reuse is to reduce their structural complexity to make them easier to apprehend. In particular, we have continued to work in collaboration with the University of Manchester on DistillFlow, an approach to remove the structural redundancy in workflows. Our contribution is four fold. Firstly, we identify a set of anti-patterns that contribute to the structural workflow complexity. Secondly, we design a series of refactoring transformations to replace each anti-pattern by a new semantically-equivalent pattern with less redundancy and simplified structure. Thirdly, we introduce a distilling algorithm that takes in a workflow and produces a distilled semantically-equivalent workflow [8]. Lastly, we provide an implementation of our refactoring approach (dedicated demo published [24]) that we evaluate on both the public Taverna workflows and on a private collection of workflows from the BioVel project. On going work includes extending the list of anti-patterns to be considered and identifying *good patterns*, that is, patterns which are easy to maintain and have systematically been able to be executed. This has been done in the context of the master internship of Stéphanie Kamgnia Wonkap [37]. First results obtained are promising.

5.5. Systems Biology

5.5.1. Analyzing SBGN-AF Networks Using Normal Logic Programs

A wide variety of signaling networks are available in the literature or in databases under the form of influence graphs. In order to understand the systems underlying these networks and to modify them for a medical purpose, it is necessary to understand their dynamics. Consequently, a variety of modelling techniques for these networks have been developped. In particular, it is possible to model their dynamical behavior with Boolean networks. The construction of these Boolean networks starting from influence graphs requires a parametrization of some Boolean functions. This task is most often realized by interpreting experimental results, that can be hard to obtain.

We introduced a method that allows to model any influence graph expressed in the Systems Biology Graphical Notation Activity Flow language (SBGN-AF) under the form of a Boolean network [32], [29]. The parametrization does not rely on any experimental results but on general principles that govern the dynamics of signaling networks. Together with the translation of a SBGN-AF influence graph into predicates, these general principles expressed under the form of logic rules form a first-order normal logic program (NLP) equivalent to a Boolean network. We show that the trajectories as well as the steady-state of any SBGN-AF network can be obtained by computing the orbits and the supported models of its corresponding NLP, respectively.

5.5.2. Scalable methods for analysing dynamics of automata networks

In collaboration with T. Chatain, S. Haar, S. Schwoon, and L. Jezeguel (INRIA MEXICO), we explored new techniques for computing the reachable attractors in automata networks using Petri net unfoldings [22]. Attractors of network dynamics represent the long-term behaviours of the modelled system. Their characterization is therefore crucial for understanding the response and differentiation capabilities of a dynamical system. In the scope of qualitative models of interaction networks, the computation of attractors reachable from a given state of the network faces combinatorial issues due to the state space explosion. Our new algorithm relies on Petri net unfoldings that can be used to compute a compact representation of the dynamics, in particular by exploiting the concurrency of the transitions in order to remove redundant sequences of transitions. We illustrate the applicability of the algorithm with Petri net models of cell signalling and regulation networks, Boolean and multi-valued. The proposed approach aims at being complementary to existing methods for deriving the attractors of Boolean models, while being generic since it actually applies to any safe Petri net.

In collaboration with M. Folschette, M. Magnin, O. Roux (IRCCYN, Nantes), and K. Inoue (NII, Tokyo), we developed a framework for identifying classical Boolean or discrete networks models from Proces Hitting (PH) models [10]. The PH allows to model non-deterministic cooperations between interacting components, and we have shown that the dynamics of a single PH can embed (include) the dynamics of multiple discrete networks, where transitions functions are deterministic. Hence, if a behaviour is shown impossible at the PH model, it is necessary impossible in any included discrete models. Such kind of analysis is relevant in systems biology, where the cooperations between components are often under-determined and the enumeration of all compatible discrete models is intractable: our framework allows to reason on the dynamics of a single abstract model.

Finally, a chapter summarizing the recent advances on static analysis for dynamics of large biological networks has been published as part of the *Logical Modeling of Biological Systems* handbook [30].

ANGE Project-Team

6. New Results

6.1. Highlights of the Year

In 2014, ANGE status turned from Inria team to Inria project-team. Afterwards, M. Parisot was recruited by Inria as a junior researcher.

6.2. Analysis of models in fluid mechanics

6.2.1. Well-posedness of multilayer Shallow Water-type equations

Participants: Emmanuel Audusse, Bernard Di Martino, Ethem Nayir, Yohan Penel.

The hyperbolicity of some 2-layer Shallow Water equations had been proven in [26], [23], there are many open theoretical investigations to lead about these systems. In particular, E. Nayir proved the local well-posedness of the model derived in [23] for periodic boundary conditions. Next steps will consist in extending this preliminary result to the whole space and proving the global existence of strong solutions. The existence of weak solutions will be studied from B. Di Martino's work. The hyperbolicity for N layers must also be investigated.

As for numerical aspects, the use of FRESHKISS3D will provide qualitative assessments for modelling issues (viscous tensor, source terms, variable density, interfacial velocities). It will also yield comparisons with theoretical results, in particular when the number of layers goes to infinity.

6.2.2. Non-hydrostatic models

Participants: Dena Kazerani, Jacques Sainte-Marie, Nicolas Seguin.

Together with Corentin Audiard from Univ. Pierre et Marie Curie, we investigated the structure of general non hydrostatic models for shallow water flows. This includes the Green–Naghdi equations and the model proposed by Bristeau *et al.* in [13]. D. Kazerani proved that such systems possess a symmetric structure based on the existence of an energy. The main difference with the well-known hyperbolic case is due to the presence of differential operators instead of matrices.

6.3. Modelling of complex flows

6.3.1. Dynamics of sedimentary river beds with stochastic fluctuations

Participants: Emmanuel Audusse, Philippe Ung.

We studied in [9] the behaviour of the solution of the Saint-Venant–Exner equations when a stochastic term is introduced in the model through the sediment flux. A first investigation was done considering periodic boundary conditions and the next part of this study is devoted to the case when physical ones are imposed. Our goal is to investigate the possibility to bring out a characteristic long time behaviour and to establish a relation between the injected noise and the physical parameters involved in the model. This work was achieved in collaboration with Sébastien Boyaval from Lab. Hydraulique Saint-Venant.

6.3.2. Non-hydrostatic effects

Participants: Nora Aïssiouene, Marie-Odile Bristeau, Edwige Godlewski, Dena Kazerani, Anne Mangeney, Jacques Sainte-Marie, Nicolas Seguin.

The objective is to derive a model corresponding to a depth averaged version of the incompressible Euler equations with free surface and to develop a robust numerical method for the resolution of the model.

Concerning the modelling aspect, a non-hydrostatic shallow water-type model approximating the incompressible Euler and Navier-Stokes sytems with free surface was developped and published in [13]. The closure relations are obtained by a minimal energy constraint instead of an asymptotic expansion. The model slightly differs from the well-known Green-Naghdi model and is confronted with stationary and analytical solutions of the Euler system corresponding to rotational flows.

The numerical approximation relies on a projection-correction type scheme. The hyperbolic part of the system is approximated using a kinetic finite volume solver and the correction step implies to solve an elliptic problem involving the non-hydrostatic part of the pressure.

In one dimension, the resolution of the incompressibility problem leads to solve a mixed problem where the pressure and the velocity are defined in compatible approximation spaces. This step uses a variationnal formulation of the shallow water version of the incompressibility condition.

This numerical scheme satisfies classical properties (positivity, well-balancing and consistency) and a discrete entropy inequality. Several numerical experiments are performed to confirm the relevance of our approach.

This approach will allow us to extend the numerical method in higher dimensions and to treat particular difficult cases occuring in specific geophysical situations (dry/wet interfaces).

6.3.3. Plasticity in Shallow Water equations

Participant: Nicolas Seguin.

In collaboration with Bruno Després and Clément Mifsud from Univ. Pierre et Marie Curie, we proposed in [20] a new definition of solutions for hyperbolic Friedrichs' systems in bounded domains, which follows the idea of Lions' dissipative solutions and Otto's boundary formulation for conservation laws. We proved in the classical settings existence and uniqueness. The goal of this project is to be able to incorporate nonlinear effects of plasticity in models of elasticity or overflowing in channels for shallow water flows, by adding entropy compatible constraints.

6.3.4. Management of marine energies

Participants: Cindy Guichard, Martin Parisot, Jacques Sainte-Marie, Julien Salomon.

The purpose of this project is to model floating devices (like buoys) in the context of recovering energy from water resources (seas and oceans). If the free surface flow can be handled by means of the Saint-Venant equations, the area under the buoys requires a different modelling (for example equivalence with springs) as the surface is constrained. The Archimedes' principle is also involved. Some preliminary numerical results were obtained thanks to the FRESHKISS3D code.

To go further, the optimisation of the overall process is also under consideration. Indeed, to maximise the amount of recovered energy, the bathymetry, the shape of the buoy, the number of buoys are critical parameters which must be modelled in view of industrial applications. Optimal control methods are applied to determine the best configuration depending on the devices: optimisation of the kinetic energy for water-turbines or of the potential energy for buoys.

6.4. Accurate simulations of fluid flows

6.4.1. A numerical scheme for the Saint-Venant–Exner equations

Participants: Emmanuel Audusse, Philippe Ung.

After having established a Godunov-type method based on the design of a three-wave Approximate Riemann Solver for the Saint-Venant equations [10], we extended this approach to the Saint-Venant–Exner equations for modelling the sediment transport. The coupled aspect between the hydraulic and the morphodynamic parts is only located on the evaluation of the wave velocities. Under this assumption, the proposed scheme can be interpreted as a hybrid method between the splitting and non-splitting methods and it also raises the issue of the choice between the two previous approaches.

These results were proven in collaboration with Christophe Chalons from Univ. Versailles–Saint-Quentin.

6.4.2. Simulations of fluid/particules interactions

Participant: Nicolas Seguin.

In collaboration with Nina Aguillon and Frédéric Lagoutière from Univ. Paris-Sud, we proved in [7] the convergence of finite volume schemes for a simplified model of fluid-particle interaction. The mesh follows the particle which appears in the model as a pointwise contribution. The numerical scheme is based on local well-balanced fluxes, which permits to obtain compactness and convergence.

6.4.3. Hydrostatic reconstruction

Participants: Emmanuel Audusse, Marie-Odile Bristeau, Jacques Sainte-Marie.

The hydrostatic reconstruction is a general and efficient method to handle source terms that uses an arbitrary solver for the homogeneous problem and leads to a consistent, well-balanced, positive scheme satisfying a semi-discrete entropy inequality.

In [8], we proved with Francois Bouchut from Univ. Marne-la-Vallée that the hydrostatic reconstruction coupled to the classical kinetic solver satisfies a fully discrete entropy inequality which involves an error term but the latter goes to zero strongly with the mesh size.

6.4.4. A numerical scheme for multilayer shallow-water model for all Froude regimes

Participant: Martin Parisot.

The aim of this work in collaboration with Jean-Paul Vila from INSA/IMT is to propose an efficient numerical resolution to simulate stratified non-miscible fluids. The strategy should be consistent for all regime especially with the so-called low-Froude regime particularly relevant for applications. The proposed scheme is entropy-satisfying, well-balanced and asymptotic preserving. In addition the stability of the scheme is ensured for large time scale. More precisely, it does not depend on the gravity waves, which are very restrictive for the targeted applications, such as oceanology and meteorology. Further work using the strategy for sustainable energies is in progress.

6.4.5. Adaptation of the Godunov scheme to the low Froude regime

Participants: Emmanuel Audusse, Do Minh Hieu, Yohan Penel.

Standard numerical schemes designed for the simulation of fluid flows are known to fail when the Mach number becomes too small. Similar behaviours are observed for geophysical flows when the Froude number decreases. Do Minh Hieu is interested in the numerical simulation of the Shallow Water equations including some Coriolis forces. He investigated several corrections of the standard Godunov schemes in 1D to preserve the kernel of spatial operators involved in the aformentioned equations and blamed for being responsible of the loss of accuracy. He now intends to perform the same analysis in 2D under the supervision of E. Audusse, S. Dellacherie (from CEA), P. Omnès (from CEA) and Y. Penel.

6.5. Software development and assessments

6.5.1. Improvements in the FRESHKISS3D code

Participants: Marie-Odile Bristeau, David Froger, Raouf Hamouda, Jacques Sainte-Marie.

Several tasks have been achieved in the FRESHKISS3D software:

- FreshKiss3D has been improved to take into account the second order in space for the 3D cases.
- The solver now includes the second order in time.
- The numerical validation using 3D numerical analytical solutions has been achieved.
- Numerous simulations have been driven by industrial contracts:
  - Simulations of fluid hydrodynamics in lagoons for optimizing the geometric field to ensure a high level of agitation for a low energy consumption (SAUR)
  - Simulations of fluid hydrodynamics in lagoons showing the vertical distribution of velocity and how to use it for optimizing micro-algae production (Salinalgue)
- Tsunamis simulations leading to the module TsunaMaths, web interface showing some historical tsunamis.
- Geometric implementations of FRESHKISS3D have been improved.
- Unit tests are being made automatically as the source code is modified.
- A user interface has been created using Python.
- The parallelization of FRESHKISS3D with MPI is under development.

ANTIQUE Team

6. New Results

6.1. Highlights of the Year

Patrick and Radhia Cousot have received in 2014 the IEEE Computer Society IEEE Computer Society Harlan D. Mills award for the invention of abstract interpretation, developpment of tool support and practical application http://www.computer.org/portal/web/awards/cousots.

6.2. Memory Abstraction

6.2.1. Modular Construction of Shape-Numeric Analyzers

Participants: Xavier Rival [correspondant], Bor-Yuh Evan Chang [University of Colorado, Boulder, USA], Huisong Li, Antoine Toubhans.

Abstract interpretation, Memory abstraction, Shape abstract domains. In [24], we discuss the modular construction of memory abstract domains.

The aim of static analysis is to infer invariants about programs that are tight enough to establish semantic properties, like the absence of run-time errors. In the last decades, several branches of the static analysis of imperative programs have made significant progress, such as in the inference of numeric invariants or the computation of data structures properties (using pointer abstractions or shape analyzers). Although simultaneous inference of shape-numeric invariants is often needed, this case is especially challenging and less well explored. Notably, simultaneous shape-numeric inference raises complex issues in the design of the static analyzer itself. We studied the modular construction of static analyzers, based on combinations of atomic abstract domains to describe several kinds of memory properties and value properties.

6.2.2. An abstract domain combinator for separately conjoining memory abstractions

Participants: Xavier Rival [correspondant], Bor-Yuh Evan Chang [University of Colorado, Boulder, USA], Antoine Toubhans.

Abstract interpretation, Memory abstraction, Shape abstract domains. In [25], we studied the separating combination of heap abstract domains.

The breadth and depth of heap properties that can be inferred by the union of today's shape analyses is quite astounding. Yet, achieving scalability while supporting a wide range of complex data structures in a generic way remains a long-standing challenge. We proposed a way to side-step this issue by defining a generic abstract domain combinator for combining memory abstractions on disjoint regions. In essence, our abstract domain construction is to the separating conjunction in separation logic as the reduced product construction is to classical, non-separating conjunction. This approach eases the design of the analysis as memory abstract domains can be re-used by applying our separating conjunction domain combinator. And more importantly, this combinator enables an analysis designer to easily create a combined domain that applies computationally-expensive abstract domains only where it is required.

6.2.3. Abstraction of Arrays Based on Non Contiguous Partitions

Participants: Xavier Rival [correspondant], Jiangchao Liu.

Abstract interpretation, Memory abstraction, Array abstract domains. In [20], we studied array abstractions.

Array partitioning analyses split arrays into contiguous partitions to infer properties of cell sets. Such analyses cannot group together non contiguous cells, even when they have similar properties. We proposed an abstract domain which utilizes semantic properties to split array cells into groups. Cells with similar properties will be packed into groups and abstracted together. Additionally, groups are not necessarily contiguous. This abstract domain allows to infer complex array invariants in a fully automatic way. Experiments on examples from the Minix 1.1 memory management demonstrated its effectiveness.

6.3. Static analysis of JavaScript applications

6.3.1. Automatic Analysis of Open Objects in Dynamic Language Programs

Participants: Arlen Cox [correspondant], Bor-Yuh Evan Chang [University of Colorado, Boulder, USA], Xavier Rival.

Abstract interpretation, Dynamically typed languages, Verification In [14], we have studied the abstraction of open objects in dynamic language programs (like JavaScript).

In dynamic languages, objects are open: they support iteration over and dynamic addition/deletion of their attributes. Open objects, because they have an unbounded number of attributes, are difficult to abstract without a priori knowledge of all or nearly all of the attributes and thus pose a significant challenge for precise static analysis. To address this challenge, we presented the HOO (Heap with Open Objects) abstraction that can precisely represent and infer properties about open-object-manipulating programs without any knowledge of specific attributes. It achieves this by building upon a relational abstract domain for sets that is used to reason about partitions of object attributes. An implementation of the resulting static analysis is used to verify specifications for dynamic language framework code that makes extensive use of open objects, thus demonstrating the effectiveness of this approach.

6.3.2. Desynchronized Multi-State Abstractions for Open Programs in Dynamic Languages

Participants: Arlen Cox [correspondant], Bor-Yuh Evan Chang [University of Colorado, Boulder, USA], Xavier Rival.

Abstract interpretation, Dynamically typed languages, Verification In [15], we have studied desynchronized multi-state abstractions for open programs in dynamic languages (libraries).

Dynamic language library developers face a challenging problem: ensuring that their libraries will behave correctly for a wide variety of client programs without having access to those client programs. This problem stems from the common use of two defining features for dynamic languages: callbacks into client code and complex manipulation of attribute names within objects. To remedy this problem, we introduced two state-spanning abstractions. To analyze callbacks, the first abstraction desynchronizes a heap, allowing partitions of the heap that may be affected by a callback to an unknown function to be frozen in the state prior to the call. To analyze object attribute manipulation, building upon an abstraction for dynamic language heaps, the second abstraction tracks attribute name/value pairs across the execution of a library. We implemented these abstractions and use them to verify modular specifications of class-, trait-, and mixin-implementing libraries.

6.4. Static analysis of Spreadsheet applications

Participants: Tie Cheng [correspondant], Xavier Rival.

Abstract interpretation, Spreadsheet applications, Verification In [13], we have proposed a static analysis to detect type unsafe operations in spreadsheet applications including formulas and macros.

Spreadsheets are widely used, yet are error-prone: they use a weak type system, allowing certain operations that will silently return unexpected results, like comparisons of integer values with string values. However, discovering these issues is hard, since data and formulas can be dynamically set, read or modified. We defined a static analysis that detects all run-time type-unsafe operations in spreadsheets. It is based on an abstract interpretation of spreadsheet applications, including spreadsheet tables, global re-evaluation and associated programs. Our implementation supports the features commonly found in real-world spreadsheets. We ran our analyzer on the EUSES Spreadsheet Corpus. This evaluation shows that our tool is able to automatically verify a large number of real spreadsheets, runs in a reasonable time and discovers complex bugs that are difficult to detect by code review or by testing.

6.5. Mechanically Verifying a Shape Analysis

Participant: Arnaud Spiwack.

Program verification, Abstract interpretation, Static analysis, Shape analysis, Coq. The result of a static analysis is only as good as the trust put into its correctness. For critical software, the standards are very high, and trusting a complex tool requires costly inspection of its implementation. Mechanically proving the correctness of static analysers is a way to lower these costs: the exigence of trust is moved from various complex dedicated tools to a single simpler general purpose one.

In this context, Arnaud Spiwack worked on an ongoing Coq implementation and certification of a shape abstract domain. The implementation, named Cosa, is based on Evan Chang and Xavier Rival's Xisa. It targets an intermediary language of Xavier Leroy's Compcert C, and interfaces with the domains of the Verasco project.

The development of Cosa lead Arnaud Spiwack to express the abstract interpretation correctness property in term of refinement calculus, which allowed to use interaction structures (a type theoretic variant of the refinement calculus) as a central structuring element of Cosa. Arnaud Spiwack started investigating how the technology of nominal sets could be leveraged to prove the correctness of unfolding (which involves choosing new names) in Cosa.

6.6. Static Analysis of Embedded Critical Concurrent Software

6.6.1. AstréeA: A Static Analyzer for Large Embedded Multi-Task Software

Participant: Antoine Miné.

In [11], we present the design, implementation and experimentation of the ASTRÉEA static analyzer, an extension of the ASTRÉE static analyzer dedicated to analyzing the run-time errors in embedded critical concurrent software. Such software are already present in critical systems and will likely become the norm with the generalization of multi-core processors in embedded systems, leading to new challenging demands in verification. One major challenge is that a concurrent program execution does not follow a fixed sequential order, but one of many interleavings of executions from different tasks chosen by the scheduler. As it is impractical to build a fully flow-sensitive analysis by enumerating explicitly all interleavings, we took inspiration from thread-modular methods: we analyze each thread individually, in an environment consisting of (an abstraction of) the effect of the other threads. This is a form of rely-guarantee reasoning, but in a fully automatic static analysis settings formalized as abstract interpretation: a thread-modular static analysis is viewed as a computable abstraction of a complete concrete, fixpoint-based thread-modular semantics. This permits a fine control between precision and efficiency, and opens the way to analysis specialization: any given safety property of a given program can be theoretically inferred given the right abstract domain. The presentation describes our subsequent work in improving the precision of ASTRÉEA by specialization on our target applications, and the interesting abstractions we developed along the way. For instance, we developed new interference abstractions enabling a limited but controllable (for efficiency) degree of relationality and flow-sensitivity. We also designed abstractions able to exploit our knowledge of the real-time scheduler used in the analysis target: i.e., it schedules tasks on a single core and obeys a strict priority scheme. The end-result is a more precise analyzer on our target applications, with currently around a thousand alarms.

6.6.2. Static Analysis by Abstract Interpretation of Concurrent Programs under the TSO Weak Memory Model

Participants: Thibault Suzanne, Antoine Miné.

In [33], we present an abstract semantics for the Total Store Ordering (TSO) memory model, a weakly consistent memory model used in major multi-core processors. This abstraction forgets some information about the order in which variables are written into by each thread. This results in a much simplified concrete semantics, but which is still not computable. We then express the semantics based on partitioned sets of points in a vector space, which allows applying classic methods from abstract interpretation (such as numeric abstract domains) to achieve a fully computable abstract semantics and automatically infer an over-approximation of the set of reachable states of a program running under the TSO memory model. The method is proved correct and, in certain cases, optimal, using the standard tools of abstraction interpretation (Galois connections).

Moreover, we have written a prototype static analyzer for simple program fragments written in an assembly-like language, and experimented our abstraction on a few small examples.

6.7. Inference of Termination and Liveness properties

6.7.1. A Decision Tree Abstract Domain for Proving Conditional Termination

Participants: Caterina Urban, Antoine Miné.

In [26], we present a new parameterized abstract domain able to refine existing numerical abstract domains with finite disjunctions. The elements of the abstract domain are decision trees where the decision nodes are labeled with linear constraints, and the leaf nodes belong to a numerical abstract domain. The abstract domain is parametric in the choice between the expressivity and the cost of the linear constraints for the decision nodes (e.g., polyhedral or octagonal constraints), and the choice of the abstract domain for the leaf nodes. We describe an instance of this domain based on piecewise-defined ranking functions for the automatic inference of sufficient preconditions for program termination. We have implemented a static analyzer for proving conditional termination of programs written in (a subset of) C and, using experimental evidence, we show that it performs well on a wide variety of benchmarks, it is competitive with the state of the art and is able to analyze programs that are out of the reach of existing methods.

6.7.2. An Abstract Domain to Infer Ordinal-Valued Ranking Functions

Participants: Caterina Urban, Antoine Miné.

The traditional method for proving program termination consists in inferring a ranking function. In many cases (i.e. programs with unbounded non-determinism), a single ranking function over natural numbers is not sufficient. In [30], we propose a new abstract domain to automatically infer ranking functions over ordinals. We extend an existing domain for piecewise-defined natural-valued ranking functions to polynomials in ω, where the polynomial coefficients are natural-valued functions of the program variables. The abstract domain is parametric in the choice of the state partitioning inducing the piecewise-definition and the type of functions used as polynomial coefficients. To our knowledge this is the first abstract domain able to reason about ordinals. Handling ordinals leads to a powerful approach for proving termination of imperative programs, which in particular allows us to take a first step in the direction of proving termination under fairness constraints and proving liveness properties of (sequential and) concurrent programs.

6.7.3. Proving Guarantee and Recurrence Temporal Properties by Abstract Interpretation

Participants: Caterina Urban, Antoine Miné.

We present in [28] a new static analysis methods for proving liveness properties of programs. In particular, with reference to the hierarchy of temporal properties proposed by Manna and Pnueli, we focus on guarantee (i.e., “something good occurs at least once”) and recurrence (i.e., “something good occurs infinitely often”) temporal properties. We generalize the abstract interpretation framework for termination presented by Cousot and Cousot. Specifically, static analyses of guarantee and recurrence temporal properties are systematically derived by abstraction of the program operational trace semantics. These methods automatically infer sufficient preconditions for the temporal properties by reusing existing numerical abstract domains based on piecewise-defined ranking functions. We augment these abstract domains with new abstract operators, including a dual widening. To illustrate the potential of the proposed methods, we have implemented a research prototype static analyzer, for programs written in a C-like syntax, that yielded interesting preliminary results.

6.8. Numeric Invariant Inference

6.8.1. A Numeric Abstract Domain to Infer Octagonal Constraints with Absolute Value

Participants: Liqian Chen [National Laboratory for Parallel and Distributed Processing, National University of Defense Technology, Changsha, P.R.China], Jiangchao Liu, Antoine Miné, Deepak Kapur [University of New Mexico, USA], Ji Wang [National Laboratory for Parallel and Distributed Processing, National University of Defense Technology, Changsha, P.R.China].

The octagon abstract domain, devoted to discovering octagonal constraints (also called Unit Two Variable Per Inequality or UTVPI constraints) of a program, is one of the most commonly used numerical abstractions in practice, due to its quadratic memory complexity and cubic time complexity. However, the octagon domain itself is restricted to express convex sets and has limitations in handling non-convex properties which are sometimes required for proving some numerical properties in a program. In [12], we intend to extend the octagon abstract domain with absolute value, to infer certain non-convex properties by exploiting the absolute value function. More precisely, the new domain can infer relations of the form $\{ \pm X \pm Y \leq c,\ \pm X \pm |Y| \leq d,\ \pm|X| \pm |Y| \leq e \}$. We provide algorithms for domain operations such that the new domain still enjoys the same asymptotic complexity as the octagon domain. Moreover, we present an approach to support strict inequalities over rational or real-valued variables in this domain, which also fits for the octagon domain. Experimental results of our prototype are encouraging; The new domain is scalable and able to find non-convex invariants of interest in practice but without too much overhead (compared with that using octagons).

6.8.2. A Method to Infer Inductive Numeric Invariants Inspired from Constraint Programming.

Participant: Antoine Miné.

In [29], we suggest the idea of using algorithms inspired by Constraint Programming in order to infer inductive invariants on numeric programs. Similarly to Constraint Programming solvers on continuous domains, our algorithm approximates the problem from above, using decreasing iterations that may split, discard, and tighten axis-aligned boxes. If successful, the algorithm outputs a set of boxes that includes the initial states and is a post-fixpoint of the abstract semantic function of interest. Our work is very preliminary; many improvements still need to be performed to determine if the method is usable in practice, and in which contexts. Nevertheless, we show that a naive proof-of-concept implementation of our algorithm is already capable of inferring non-trivial inductive invariants that would otherwise require the use of relational or even non-linear abstract domains when using more traditional abstract interpretation iteration methods.

6.9. Bisimulation Metrics

6.9.1. Bisimulation for Markov Decision Processes through Families of Functional Expressions

Participants: Norman Ferns, Sophia Knight [LIX, France], Doina Precup [McGill University, Canada].

Markov decision processes, Bisimulation, Metrics.

In [17], we have transfered a notion of quantitative bisimilarity for labelled Markov processes [51] to Markov decision processes with continuous state spaces. This notion takes the form of a pseudometric on the system states, cast in terms of the equivalence of a family of functional expressions evaluated on those states and interpreted as a real-valued modal logic. Our proof amounts to a slight modification of previous techniques [56], [55] used to prove equivalence with a fixed-point pseudometric on the state-space of a labelled Markov process and making heavy use of the Kantorovich probability metric. Indeed, we again demonstrate equivalence with a fixed-point pseudometric defined on Markov decision processes [52]; what is novel is that we recast this proof in terms of integral probability metrics [54] defined through the family of functional expressions, shifting emphasis back to properties of such families. The hope is that a judicious choice of family might lead to something more computationally tractable than bisimilarity whilst maintaining its pleasing theoretical guarantees. Moreover, we use a trick from descriptive set theory to extend our results to MDPs with bounded measurable reward functions, dropping a previous continuity constraint on rewards and Markov kernels.

6.9.2. Bisimulation Metrics are Optimal Value Functions

Participants: Norman Ferns, Doina Precup [McGill University, Canada].

Markov decision processes, Bisimulation, Metrics.

Bisimulation is a notion of behavioural equivalence on the states of a transition system. Its definition has been extended to Markov decision processes, where it can be used to aggregate states. A bisimulation metric is a quantitative analog of bisimulation that measures how similar states are from a the perspective of long-term behavior. Bisimulation metrics have been used to establish approximation bounds for state aggregation and other forms of value function approximation. In [18], we prove that a bisimulation metric defined on the state space of a Markov decision process is the optimal value function of an optimal coupling of two copies of the original model. We prove the result in the general case of continuous state spaces. This result has important implications in understanding the complexity of computing such metrics, and opens up the possibility of more efficient computational methods.

6.10. Abstraction of Rule-Based Biological Models

6.10.1. Stochastic fragments: A framework for the exact reduction of the stochastic semantics of rule-based models

Participants: Jérôme Feret, Heinz Koeppl [École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland], Tatjana Petrov [École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland].

Protein-protein interaction networks, Stochastic systems, Backward bisimulations, Model reduction. In [9], we propose an abstract interpretation-based framework for reducing the state space of stochastic semantics for protein-protein interaction networks. Our approach consists in quotienting the state space of networks. Yet interestingly, we do not apply the widely-used strong lumpability criterion which imposes that two equivalent states behave similarly with respect to the quotient, but a weak version of it. More precisely, our framework detects and proves some invariants about the dynamics of the system: indeed the quotient of the state space is such that the probability of being in a given state knowing that this state is in a given equivalence class, is an invariant of the semantics. Then we introduce an individual-based stochastic semantics (where each agent is identified by a unique identifier) for the programs of a rule-based language (namely Kappa) and we use our abstraction framework for deriving a sound population-based semantics and a sound fragments-based semantics, which give the distribution of the traces respectively for the number of instances of molecular species and for the number of instances of partially defined molecular species. These partially defined species are chosen automatically thanks to a dependency analysis which is also described in [9].

6.10.2. An algebraic approach for inferring and using symmetries in rule-based models

Participant: Jérôme Feret.

Graph rewriting, Single-pushout semantics, Symmetries, Bisimulations, Model reduction. Symmetries arise naturally in rule-based models, and under various forms. Besides automorphisms between site graphs, which are usually built within the semantics, symmetries can take the form of pairs of sites having the same capabilities of interactions, of some protein variants behaving exactly the same way, or of some linear, planar, or 3D molecular complexes which could be seen modulo permutations of their axis and/or mirror-image symmetries. In [16], we propose a unifying handling of symmetries in Kappa. We follow an algebraic approach, that is based on the single pushout semantics of Kappa. We model classes of symmetries as finite groups of transformations between site graphs, which are compatible with the notion of embedding (that is to say that it is always possible to restrict a symmetry that is applied with the image of an embedding to the domain of this embedding) and we provide some assumptions that ensure that symmetries are compatible with pushouts. Then, we characterise when a set of rules is symmetric with respect to a group of symmetries and, in such a case, we give sufficient conditions so that this group of symmetries induces a forward bisimulation and/or a backward bisimulation over the population semantics.

6.11. Model checking of Logical Biological Models

6.11.1. Model checking logical regulatory networks

Participants: Pedro T. Monteiro [INESC-ID, Lisboa, Portugal], Wassim Abou-Jaoudé, Denis Thieffry [IBENS, France], Claudine Chaouiya [IGC, Oeiras, Portugal].

Model checking, Regulatory networks. Regulatory and signalling networks control cell behaviours in response to environmental cues. The logical formalism has been widely employed to study these interaction networks, which are modelled as discrete dynamical systems. While biologists identify networks encompassing more and more components, properties of biological relevance become hard to verify.

In [22], we report on the use of model-checking techniques to address this challenge. This approach is illustrated by an application dealing with the modelling of T-helper lymphocyte differentiation.

6.11.2. Model checking to assess T-helper cell plasticity

Participants: Wassim Abou-Jaoudé, Pedro T. Monteiro [INESC-ID, Lisboa, Portugal], Aurélien Naldi [Centre Intégratif de Lausanne, Lausanne, Switzerland], Maximilien Grandclaudon [Institut Curie, Paris, France], Vassili Sommeils [Institut Curie, Paris, France], Claudine Chaouiya [IGC, Oeiras, Portugal], Denis Thieffry [IBENS, France].

Model checking, Logical modeling. Computational modeling constitutes a crucial step toward the functional understanding of complex cellular networks. In particular, logical modeling has proven suitable for the dynamical analysis of large signaling and transcriptional regulatory networks. In this context, signaling input components are generally meant to convey external stimuli, or environmental cues. In response to such external signals, cells acquire specific gene expression patterns modeled in terms of attractors (e.g., stable states).The capacity for cells to alter or reprogram their differentiated states upon changes in environmental conditions is referred to as cell plasticity. In this article, we present a multivalued logical framework along with computational methods recently developed to efficiently analyze large models.We mainly focus on a symbolic model checking approach to investigate switches between attractors subsequent to changes of input conditions. As a case study, we consider the cellular network regulating the differentiation of T-helper (Th) cells, which orchestrate many physiological and pathological immune responses. To account for novel cellular subtypes, we present, in [8], an extended version of a published model of Th cell differentiation.We then use symbolic model checking to analyze reachability properties between Th subtypes upon changes of environmental cues. This allows for the construction of a synthetic view of Th cell plasticity in terms of a graph connecting subtypes with arcs labeled by input conditions. Finally, we explore novel strategies enabling specific Th cell polarizing or reprograming events.

AOSTE Project-Team

6. New Results

6.1. Languages, Models of Computation and Metamodeling using logical clock constraints

Participants: Julien Deantoni, Robert de Simone, Frédéric Mallet, Marie Agnès Peraldi Frati.

A revised and updated version of our previous work on UML MARTE Time Model was written in survey textbook form for a larger audience, and published in [38]. Same was done for the more applied specific findings of the ARTEMIS PRESTO European project [39]. Also, a research report finalizing the denotational semantics of the logical clock constraint languages was issued for reference [44].

6.2. Experiments with Architecture and Application modeling

Participants: Robert de Simone, Émilien Kofman, Jean-Vivien Millo, Amine Oueslati, Mohamed Bergach.

We submitted for publication our theoretical results on formal mapping of an application written as a process network dataflow graph onto an abstract architecture model involving a network-on-chip and manycore processor arrays [24].

In the context of the *FUI Clistine* collaborative project (which aims at building a cheap supercomputer by assembling low-cost, general-purpose and network processors interconnected by a time-predictable, on-board network), we considered the issue of classifying general application types, in the fashion inherited from UC. Berkeley's 13 "dwarfs" [46]. Meanwhile,the modeling of desired architecture was slightly postponed due to hesitations from the main industrial partner (that will build the prototype itself). This work was the topic of Amine Oueslati's first year PhD. The classification, and the use of distinct type properties for efficient and natural encoding, was applied on typical application programs provided by partners (Galerkin methods for electromagnetic simulation by the Nachos Inria team, ray-tracing algorithms by the Optis/Simplysim SME design company).

In the context of Mohammed Bergach's CIFRE PhD contract with Kontron Toulon, we conducted an advanced modeling exercice on how to best fit large DFT (Discrete Fourier Transform) modules onto a specific processor architecture (first Intel Sandybridge, then Haswell) that offers computing compromise costs (in performance vs power) between regular CPUs and GPU hardware accelerators. There were two issues: first, how to best dimension the size of the largest FFT block that may be performed locally on a corresponding GPU compute block; second, how to distribute the many such optimal size FFT block needed in a typical radar application, using the GPU and CPU features at the best of their capacity, with account of the slow data transfer latencies across memory banks (to and from the GPU registers).
As a side-effect, people from Kontron are now using and distributing to their customers the FFT GPU libraries with ad-hoc FFT variants matching the GPU block memory sizes. The development, rather lenghty in the case of Sandybridge, was quickly adjusted and ported for Haswell. A new workshop paper is under submission.

6.3. Multiview modeling with performance and power aspects

Participants: Julien Deantoni, Ameni Khecharem, Robert de Simone, Emilien Kofman, Carlos Gomez.

In the context of the ANR HOPE project and The CIM PACA Design Platform, we continued our work on joint modeling and co-simulation of abstract architecture and application (use case scenario) models, together with non-functional aspect views such as performance, power and temperature. The goal of the HOPE project is to consider *hierarchical* {Power/Performance/Temperature} Management Units (MU), and our target is to connect our IDM modeling with dedicated tools such as Synopsys Platform Architect or Docea Power AcePlorer. The IP_Xact interface format for IP blocks is also aimed for compositional assembly representation, including non-functional properties and timing semantics constraints for co-simulation. This work is mostly continued from the former PhD thesis of Carlos Gomez to a new framework by Ameni Khecharem, as part of her PhD. Practical co-simulation trends are also investigated. Currents results were reported in [29]

6.4. Heterogeneous Languages Coordination with Concurrency and Time

Participants: Julien Deantoni, Matias Vara Larsen, Robert de Simone, Frédéric Mallet.

In the context of the ANR GEMOC project and in closely related to the mutiview approach of the team, we focused on how to deal with analysis and simulation of heterogeneous languages. Supporting coordinated use of heterogeneous domain specific languages leads to what we called the globalization of modeling language [22]. Concretely, we proposed to define a language behavioral interface to exhibit the concurrency and time aspects of the semantics of a language. The concurrency and time aspects are described by a formal extension of CCSL, named MoCCML (Model of Concurrency and Communication Modeling Language [45], [28]). Any models that conform such language exhibit a symbolic representation of all its acceptable schedules. Based on this, we shown that it is possible to coordinate heterogeneous models .To avoid redundant model coordination activities, we reified the know-how about model coordination in BCOoL (Behavioral Coordination Operator Language[34]), a language dedicated to language coordination. This work is mainly realized by Matias Vara Larsen, as part of his PhD. In this context, we organize the community around such subject for the second year in an international workshop [43] with an increasing number of participants.

6.5. Performance study of Massively Parallel Processor Array (MPPA) SoC architecture

Participants: Sid Touati, Franco Pestarini.

From a previous collaboration programme, we (Aoste Sophia) possess a MMPA manycore chip, designed and produced by the company Kalray, in Grenoble. The chip integrates 256 cores, composed of 16 clusters (themselves each with 16 cores), and a powerfull network-on-chip interconnect mesh structure. This architecture is oriented towards high performance embedded application, with real time constraints. The cores and NoC were designed to deliver predictable performance.

Our current project, during Franco Pestarini Inria International Intern period, was to test the performance of the NoC, trying to obtain better knowledge of its behavior. We put up a set of microbenchmarks to exercice the network under different specific scenarios (low overhead network traffic, high traffic), and analyzed the experimental results.

We produced a detailed deliverable report explaining under which conditions the NoC could deliver stable and predictable performances. We identified potential configurations where the network becomes unstable (leading to variable and impredictable performances and bandwith).

Meanwhile, the textbook on low-level code optimization, written between Sid Touati and Kalray CTO, appeared in published form [42]. Its content reports on some of the techniques used inside the MPPA compilation environment, and beyond.

6.6. Parametric and Non Parametric Statistics for Code Performance Analysis

Participant: Sid Touati.

This activity is conducted by Sid Touati in collaboration with Julien Worms, an associate professor in Mathematics at the university of Versailles Saint Quentin. It was started under the consideration that the performances of programs are hardly ever represented by a gaussian distribution. So, our purpose here is to study parametric statistics for analyzing the performances of programs. We are interested in modelling program performances by gaussian mixtures (using mixmod method). After a statistical modeling, we deduce multiple performance tests and performance criteria to decide with a high degree of confidence about the "best" program run version. This is still work-in-progress: we are implementing a free software for analysis based on our approach, and we are writing a rather complete research report prior to further publications in conferences and journals.

6.7. Uniprocessor Real-Time Scheduling

Participants: Falou Ndoye, Yves Sorel, Walid Talaboulma.

6.7.1. Real-Time Scheduling with Exact Preemption Cost

Previous years, we worked on schedulability analyses of dependent tasks, executed on a uniprocessor, which take into account the exact preemption cost and more generally the cost of the real-time operating system. Indeed, this cost is composed of the cost of the preemptions and the cost of the scheduler. Our approach is based on an offline real-time schedulability analysis, proved sustainable, that produces a scheduling table. This latter contains the next instants (activation and completion of tasks) when the scheduler will be called, beeing aware of the instants where tasks are preempted and then resumed. This approach allows the schedulability analysis to account preemption costs involved by other preemptions. The scheduling table contains also the address of the next task to execute preventing the scheduler to choose it in the ready tasks list, unlike with classical on-line scheduler. The theoretical results in the uniprocessor case, are given in the Falou Ndoye's PhD thesis [19] defended in April this year. This approach has been implemented through an offline scheduler that is triggered by a timer when this latter is equal to zero and loaded with the next instant contained in the scheduling table, according to a time trigger approach.

We carried out two kinds of implementations. Actually, the first one is a simulation since our time trigger offline scheduler is modelled as a high priority task running uppon an existing operating system. We experimented this approach with Vanilla Linux (not modified) and Linux/Xenomai, real-time versions of the Linux operating system, with low latency caracteristics. Of course, these implementations were only able to show that the theoterical results were correct, but did not provide good real-time performances nor a robust way to measure time without influencing the usefull code. Therefore, we implemented our scheduler on a bare metal ARM968E-S processor based on an ARM9 architecture since it is widespread in the industry world, and we experimented this processor few years ago to determine the cost of classical online schedulers.

For this purpose we used a MCB2929 developpement board, from Keil, containing the LPC2929 SoC including the ARM968E-S, an accurate timer, and various peripherals. The scheduling table is generated offline for a set of tasks, and stored in the memory as an array of couples, each composed of the task to execute and the duration elapsing until the next scheduler call. This duration is used to set the timer counter. When it hits zero it triggers a high priority interruption that is serviced by calling again the scheduler to choose the next couple (task, duration), and so on up to the end of the scheduling table. This will repeat infinitely from the beginning of the scheduling table.

We tested different set of tasks with multiple preemption scenarios, that can yield to deadlines misses. We measured for a **12Mhz** CPU and timer clock frequencies a value of **28** μ**s** for the scheduler cost, and of **1** μ**s** for the preemption cost.

6.8. Multiprocessor Real-Time Scheduling

Participants: Aderraouf Benyahia, Laurent George, Falou Ndoye, Dumitru Potop-Butucaru, Yves Sorel, Meriem Zidouni.

6.8.1. Multiprocessor Partitioned Scheduling with Exact Preemption Cost

Since we chose a multiprocessor partitioned scheduling approach, we can take advantage of the results we obtained in the case of uniprocessor real-time Scheduling accounting for the cost of the real-time operating system, i.e. the cost of preemptions and of the scheduler. From the point of view of the off-line real-time schedulability analysis we only have to consider in addition to activation and completion instants, reception of data instants. This latter instant is determinated by supposing that the cost of every data transfer is known for every possible communication medium. Indeed, when two dependent tasks are allocated to two different processors, the consuming task will have to wait for the data sent by the producing task. The theoretical results in the multiprocessor case, are given in the Falou Ndoye's PhD thesis [19] defended in April this year. We chose the message passing protocol for interprocessor communications achieved through a switched ethernet network. In order to determine precisely the cost of data transfers, we started to investigate the possible approaches to synchronize the send and receive tasks located in two different processors and to schedule them with the other tasks allocated to the same processor. This synchronization protocol will be taken into account to determine the interprocessor communication costs. Concerning these communication costs, we consider FIFO and FIFO* schedulings in the switches, the later is a FIFO scheduling based on the release time of frames at their source node. We have first corrected the trajectory approach (recently shown to be optimistic for corner cases) with FIFO scheduling to compute worst case end-to-end communication costs. Then, we have extended the trajectory approach to FIFO* scheduling. We want to implement our off-line scheduler on every processor of a multiprocessor architecture composed of at least three processors, communicating through an ethernet switch.

Concerning the delay of communications, we consider FIFO and FIFO* schedulings in the switches, the later is a FIFO scheduling based on the release time of frames at their source node. We have first corrected the trajectory approach (recently shown to be optimistic for corner cases) with FIFO scheduling to computed worst case end-to-end communication delays. Then, we have extended the trajectory approach to FIFO* scheduling.

6.8.2. Mutiprocessor Parallel Directed Acyclic Graph (DAG) scheduling

We are interested in studying the hard real-time scheduling problem of parallel Directed Acyclic Graph (DAG) tasks on multiprocessor systems. In this model, a task is defined as a set of dependent subtasks that execute under precedence constraints. The execution order of these subtasks is dynamic, i.e., a subtask can execute either sequentially or in parallel with its siblings based on the decisions of the real-time scheduler. To this end, we analyze two DAG scheduling approaches to determine the execution order of subtasks: the Model Transformation and the Direct Scheduling approaches. We consider global preemptive multiprocessor scheduling algorithms to be used with the scheduling approaches, such as Earliest Deadline First (EDF) and Deadline Monotonic (DM).

6.8.3. Gateway with Modeling Languages for Certified Code Generation

This work was carried out in the P FUI project 8.2.2 We continued the work on the gateway between the P formalism and SynDEx, started the last two years. We have integrated in the gateway the IF and FOR blocks of Simulink that were missing in the functional specification, except for particular cases where the IF block is nested in the FOR block, or the opposite. The integration of the MERGE and MUX blocks are still to be done. We extended the P formalism with architectural elements that SynDEx needs to perform schedulability analyses on functional specifications. These architectural elements are hardware resources (processor, bus, shared memory, router) and timing characteristics (deadline, period, WCET, WCTT). We developed a new part in the gateway which transforms an architectural model described with the P formalism in the input format of SynDEx. We developed also a third part in the gateway which feedbacks the schedulability analysis results obtained with SynDEx (the scheduling table) and stores them into models described with the P formalism. Finally, we have collaborated with the industrial partners to test our gateway on their use cases.

6.8.4. SynDEx updates

The first tests on the alpha version of SynDEx V8, released last year, shown some bugs that we fixed. This first release did not include a code generator. Thus, we worked to interface the distributed real-time embedded code generator of SynDEx V7 with SynDEx V8.

6.9. Probabilistic Real-Time Systems

Participants: Liliana Cucu-Grosjean, Robert Davis, Adriana Gogonel, Codé Lo, Dorin Maxim, Cristian Maxim.

The advent of complex hardware, in response to the increasing demand for computing power in next generation systems, exacerbates some of the limitations of static timing analysis for the estimation of the worst-case execution time (WCET) estimation. In particular, the effort of acquiring (1) detail information on the hardware to develop an accurate model of its execution latency as well as (2) knowledge of the timing behaviour of the program in the presence of varying hardware conditions, such as those dependent on the history of previously executed instructions. These problems are also known as the timing analysis walls. The probabilistic timing analysis, a novel approach to the analysis of the timing behaviour of next-generation real-time embedded systems, provides answers to timing analysis walls. In [23] we have described the vision of FP7 IP PROXIMA, project that is interested in the introduction of randomization of the architectures at cache level. For this type of architecture static probabilistic timing analysis is possible [20] by providing bounds on the probabilistic execution time of a task. An industrial case study from avionics is detailed in [32]. Such distribution is then used as input for probabilistic scheduling as described in [37], [36].

This year we have also provided a complete state of the art of the probabilistic real-time systems in [17].

6.10. Off-line (static) mapping and WCET analysis of real-time applications onto NoC-based many-cores

Participants: Dumitru Potop Butucaru, Thomas Carle, Manel Djemal, Robert de Simone, Zhen Zhang.

Modern computer architectures are increasingly relying on multi-processor systems-on-chips (MPSoCs, also called chip-multiprocessors), with data transfers between cores and RAM banks managed by on-chip networks (NoCs). This reflects in part a convergence between embedded, general-purpose PC, and high-performance computing (HPC) architecture designs.

In past years we have identified and compared the hardware mechanisms supporting precise timing analysis and efficient resource allocation in existing NoCs. We have determined that the NoC should ideally provide the means of enforcing a global communications schedule that is computed off-line and which is synchronized with the scheduling of computations on CPU cores. Furthermore, if in addition the computation and memory resources of the MPSoC have support for real-time predictability, then parallel applications can be developed that allow very precise WCET analysis of parallel code. WCET analysis of parallel code is joint work with Isabelle Puaut of Inria, EPI ALF.

This year we have completed our mapping (allocation and scheduling) and code generation technique and tool for NoC-based MPSoCs. NoCs pose significant challenges to both on-line (dynamic) and off-line (static) real-time scheduling approaches [25]. They have large numbers of potential contention points, have limited internal buffering capabilities, and network control operates at the scale of small data packets. Therefore, efficient resource allocation requires scalable algorithms working on hardware models with a level of detail that is unprecedented in real-time scheduling.

We considered a static (off-line) scheduling approach, and we targeted massively parallel processor arrays (MPPAs), which are MPSoCs with large numbers (hundreds) of processing cores. We proposed a novel allocation and scheduling method capable of synthesizing such global computation and communication schedules covering all the execution, communication, and memory resources in an MPPA. To allow an efficient use of the hardware resources, our method takes into account the specificities of MPPA hardware and implements advanced scheduling techniques such as pre-computed preemption of data transmissions[26] and pipelined scheduling [21].

Our method has been implemented within the Lopht tool presented in section 5.4 , and first results are presented in [26], [25], and in extenso in the PhD thesis of manel Djemal [18]. One of the objectives of the starting CAPACITES project is the evaluation of the possibility of porting Lopht and the WCET analysis technique for parallel code onto the Kalray MPPA platform.

6.11. Real-time scheduling and code generation for time-triggered platforms

Participants: Dumitru Potop Butucaru, Thomas Carle, Raul Gorcitz, Yves Sorel.

We have continued this year the work on real-time scheduling and code generation for time-triggered platforms. Much of this work was carried out as part of a bilateral collaboration with Airbus DS and the CNES, which fund the post-doctorate of Raul Gorcitz, and in our collaboration with the IRT SystemX, project FSF.

The objective is to facilitate the development of complex time-triggered systems by automating the allocation, scheduling, and code generation steps. We show that full automation is possible while taking into account all the specification elements required by a complex, real-life embedded control system. The main originality of our work is that it takes into account at the same time multiple complexity elements: functional specifications with conditional execution and multiple modes and various types of non-functional properties: real-time (release dates, deadlines, major time frame, end-to-end flows), ARINC 653 partitioning (which we can fully or partially synthesize), task preemptability, allocation. Our algorithms allow the automatic allocation and scheduling onto multi-processor (distributed) systems with a global time base, taking into account communication costs.

While the past years were mainly dedicated to the development of this scheduling and code generation technique, this year the technique and the associated tool have matured enough to allow the publication of the first results concerning the optimized scheduling algorithms [21] and its application on large case studies. Ongoing work by the post-doc Raul Gorcitz, funded by Airbus DS and the CNES aims at evaluating the applicability of our methods on embedded platforms that are being considered for the future european space launchers. The Lopht tool is also used in the IRT SystemX, project FSF as part of the proposed design flow. All extensions have been implemented in the Lopht tool. All this work has been presented *in extenso* in the PhD thesis of Thomas Carle [16].

APICS Project-Team

6. New Results

6.1. Source recovery problems

Participants: Laurent Baratchart, Sylvain Chevillard, Juliette Leblond, Christos Papageorgakis, Olga Permiakova, Dmitry Ponomarev.

The research in this section is partly joint work with Qian Tao (Univ. Macao).

It was proved in [38] that a vector field with $n+1$ components on $\mathbb{R}^n$ can be expressed uniquely as the sum of (the trace on $\mathbb{R}^n$ of) a harmonic gradient in the upper half-space, of (the trace on $\mathbb{R}^n$ of) a harmonic gradient in the lower half-space, and of a tangential divergence free vector field on $\mathbb{R}^n$. This decomposition, that we call the *Hardy-Hodge* decomposition, is valid not only for L^p vector fields as mentioned in Section 3.3.1 , but in much more general distribution spaces like $W^{-\infty,p}$ which contains all distributions with compact support or $BMO^{-\infty}$ which contains all finite sums of derivatives of bounded functions. This year we extended the decomposition to smooth hypersurfaces, where divergence-free distributions may be defined as those annihilating tangential gradient vector fields. We also studied the case where the hypersurface is only Lipschitz smooth, and then we proved the decomposition in L^p provided that p is close enough to 2 (how close depends on the Lipschitz constant of the hypersurface).

The Hardy-Hodge decomposition was used in [38] to find the kernel of the planar magnetization operator, namely a potential of the form (1) with m supported in a plane generates the zero field above that plane if, and only if there is no harmonic gradient from below in the Hardy-Hodge decomposition of m. The above mentioned generalization is now to the effect that a magnetization supported on a bounded closed surface (*e.g.* a sphere) is silent in the unbounded component of the complement of that surface if, and only if there is no harmonic gradient from inside in its Hardy-Hodge decomposition. An article is being written on this topic.

We also considered the case where m is compactly supported in the bounded component of the complement of that surface. Then m is silent if and only if it is the sum of a divergence-free distribution and of finitely many derivatives of gradients of Sobolev functions having zero trace on the surface [41].

These results shed light on the indeterminacy of inverse source problems.

6.1.1. EEG

This work is conducted in collaboration with Maureen Clerc and Théo Papadopoulo from the Athena EPI, and with Jean-Paul Marmorat (Centre de mathématiques appliquées - CMA, École des Mines de Paris).

In 3-D, functional or clinically active regions in the cortex are often modeled by point-wise sources that must be localized from measurements of a potential on the scalp. Inside the cortex, identified to a ball after the cortical mapping step, the potential satisfies a Poisson equation whose right-hand side is a linear combination of gradients of Dirac masses (the sources in EEG). In the work [3] it was shown how best rational approximation on a family of circles, cut along parallel planes on the sphere, can be used to recover the sources when they are at most 2 of them. Later, results on the behavior of poles in best rational approximation of fixed degree to functions with branch points [6] helped justifying the technique for finitely many sources (see section 4.2).

The dedicated software FindSources3D (see section 5.6), developed, in collaboration with the team Athena and the CMA, dwells on these ideas. Functions to be approximated in 2-D slices turn out to have additional *multiple* poles at their branch points so that, in the rational approximation step, it is beneficial to consider approximants with multiple poles as well (for EEG data, one should consider *triple* poles). Though numerically observed in [9], there is no mathematical justification so far why these multiple poles are attracted more strongly than simple poles to the singularities of the approximated function. This intriguing property, however, definitely helps source recovery [28]. This year we used it to automatically estimate the “most plausible” number of sources (numerically: up to 3 at the moment). Such enhancements were prompted by a developing collaboration with the BESA company, which is interested in automatic detection of the number of sources (which was left to the user until recently).

Soon, magnetic data from MEG (magneto-encephalography) will become available together with EEG data; indeed, it is now possible to use simultaneously the corresponding measurement devices. We expect this to improve the accuracy of our algorithms.

In relation to other brain exploration modalities like electrical impedance tomography (EIT, see [16]), we also consider identifying electrical conductivity in the head. This is the topic of the PhD of C. Papageorgakis, co-advised with the Athena project-team and BESA GmbH. Specifically, in layered models, we are concerned with estimating conductivity of the skull (intermediate layer). Indeed, the skull consists of a hard bone part, the conductivity of which is more or less known, and spongy bone compartments whose conductivities may vary considerably with individuals.

A preliminary question in this connection is: can one uniquely recover a homogeneous skull conductivity from a single EEG recording when the sources and the conductivities of other layers are known? And if sources are not known, which additional information do we need? These are issues currently under investigation. To put them into perspective, recall the famous Caldèron problem of deducing a bounded (nonconstant) conductivity from the knowledge of all possible pairs consisting of a potential and its current flux at the boundary. In dimension 3, when the conductivity is not smooth (less than $3/2$ of a derivative), it is unknown whether the problem is even injective (*i.e.* if two conductivities can have the same pairs of boundary potential and flux). A weaker, discrete version of this problem is: if the conductivity takes on finitely many values and the geometry of the level sets is known, does a finite set of pairs of boundary potential and flux allow one to recover it? This is a significant question to be tackled for the purpose of source recovery in EEG with known geometry but unknown conductivities inside the head.

6.1.2. Inverse Magnetization problems

This work is carried out in the framework of the "équipe associée Inria" IMPINGE, comprising Eduardo Andrade Lima and Benjamin Weiss from the Earth Sciences department at MIT (Boston, USA) and Douglas Hardin and Edward Saff from the Mathematics department at Vanderbilt University (Nashville, USA),

Localizing magnetic sources from measurements of the magnetic field away from the support of the magnetization is the fundamental issue under investigation by IMPINGE. The goal is to determine magnetic properties of rock samples (*e.g.* meteorites or stalactites), from fine field measurements close to the sample that can nowadays be obtained using SQUIDs (superconducting coil devices). Currently, rock samples are cut into thin slabs and the magnetization distribution is considered to lie in a plane, which makes for a somewhat less indeterminate framework than EEG because "less" magnetizations can produce the same field (for the slab has no inner volume). Note however that EEG data consist of both potential and current values at the boundary, whereas in the present setting only values of the normal magnetic field are provided to us.

Figure 5 presents a schematic view of the experimental setup: the sample lie on a horizontal plane at height 0 and its support is included in a rectangle. The vertical component B_z of the field produced by the sample is measured on points of a horizontal $N \times N$ rectangular grid at height h.

We set up last year a heuristic procedure to recover regularly spaced dipolar magnetizations, *i.e.* magnetizations composed of dipoles placed at the points of a regular rectangular $n \times n$ grid. The latter seems general enough a model class to approximate magnetizations commonly encountered in samples. However, for reasons of computational complexity, n is significantly smaller than N which limits the power of the model. Each dipole of the $n \times n$ grid is determined by the 3 components of its moment, thus the magnetization can be represented by a real $3n^2$-vector. If we denote by A the matrix of the operator that maps such a vector X to the vector b of measurements (which belongs to $\mathbb{R}^{N^2}$), we want to find X such that AX is close to b. For computational simplicity, we use a Euclidean criterion $\|AX - b\|_2$, which reduces the problem to a singular value decomposition of A. The inverse problem being ill-posed, A is poorly conditioned and we must resort to a regularization technique. The one we developed initially has been based on iteratively cropping the support of b, using a threshold on the intensity of the dipoles at each step, so as to reduce the number of active components

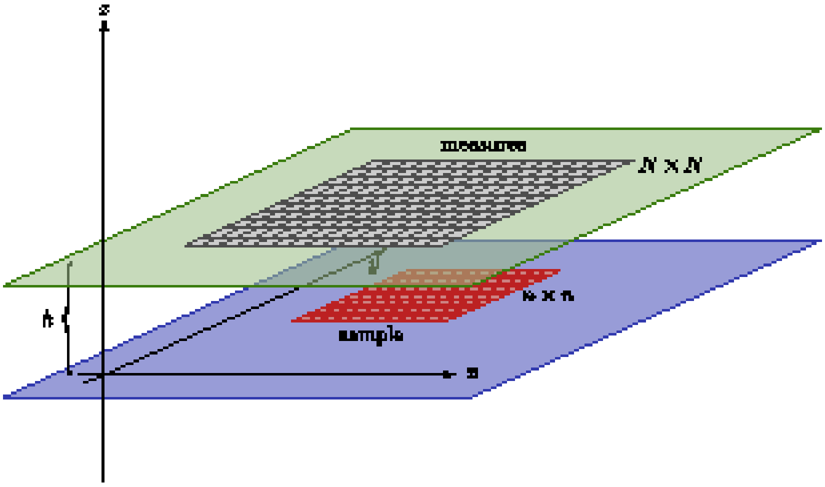


Figure 5. Schematic view of the experimental setup

in b. Preliminary experiments were performed last year on synthetic data and also on a real example (Lonar spherule).

This year, we performed more systematic experiments on real data (namely Allende chondrules and Hawaian basalt) provided by the SQUID scanning microscope at MIT lab. Cropping the support of b using thresholding has proved efficient to improve ill-conditioning for samples with localized support embedded in the slab (*e.g.*, chondrules). On the other hand, when the support of the sample is spread out (*e.g.*, Hawaian basalt), the reduction of active components of b was insignificant. We used this inversion procedure to estimate the net moment. The importance of the latter has been emphasized by the geophysicists at MIT for at least two reasons: firstly it yields important geological information on the sample in particular to estimate the magnitude of the ambient magnetic field at the time the rock was formed. Secondly, it may to some extent be measured independently, using a magnetometer, thereby allowing one to cross-validate the approach. A third, computational reason is that knowledge of the net moment should pave the way to a numerically stable reconstruction of an equivalent unidirectional magnetization. The support of the latter would provide us with valuable information to test for unidirectionality of the true magnetization, which is an important question to physicists.

When the support can be significantly shrunk while keeping the residue small (*i.e.*, explaining the data satisfactorily), estimates of the net moment based on the dipolar model obtained by inversion seem to be good. They apparently supersede the measurements by magnetometers as well as by dipole fitting procedures set up at MIT. It is interesting to notice that the magnetization obtained by our inversion procedure, either before or after shrinking the support, often does not resemble the true magnetization, even when it yields correct moment and field. This can be seen on synthetic examples and may be surmised on real data, thereby confirming that recovering the net moment and recovering the magnetization are rather different problems, the latter being considerably more ill-posed than the former.

One specific difficulty with chondrule-type examples has been to account for their thickness: they are indeed small spheres and their 3-D character cannot be completely ignored. In order to use the inversion procedure set up in the plane, we investigated the following question. Assume that the sample has some thickness, but small enough that the magnetization at a point $P = (x, y, z)$ inside the sample depends only on x and y (possibly weighted by some function that depends only on z), *i.e.* that it is of the form $m(x, y)\phi(z)$. If we consider a (truly) planar magnetization with the same distribution $m(x, y)$ but on a plane lying at some nonzero height

ε, how to choose ε so as to produce a field at height h which is closest to the field produced by the thick magnetization? This has been the object of the internship of Olga Permiakova who used local expansion of the dipole-to-field map (see her report [0]). An article is being written on this subject.

The case where the magnetization is flat but spread out on the sample is more difficult. First of all, the computational effort becomes significant and led us to use the cluster at Inria Sophia Antipolis. We succeeded in obtaining full inversions for the Hawaian basalt. The residue (approximation error) is moderate but not impressively small, which indicates that we reach the limit of modeling magnetizations by a regular grid of dipoles. However the computation of the moment compares favorably with estimates previously obtained by a different technique at MIT lab. Still, using a cluster and two days of evaluation to obtain a coarse estimate of the net moment of a sample is rather inefficient and calls for new investigations.

We also experimented an alternative regularization procedure, based on L^2 minimization under L^1 penalty as solved by the SALSA algorithm. Such methods are quite popular today for sparse recovery. However, the computational load, as well as the quality of the results, do not differ significantly from those obtained previously.

We now develop new methods in order to estimate the net moment of the magnetization, based on improvements of previously used Fourier techniques, and recently we reformulated the problem with the help of Kelvin transforms. It has been realized that the success of net moment recovery hinges on the ability to extrapolate the measurements. In particular, we managed to considerably improve previous estimates by means of data extension based on dipolar field asymptotics.

In the course of inverting the field map, we singled out magnetizations which are numerically (almost) silent from above though not from below. This illustrates how ill-posed (unstable) the problem, as theory predicts that no compactly supported magnetization can be exactly silent from above without being also exactly silent from below. Although such magnetizations seem to have small moment and therefore do not endanger the possibility of recovering the net moment, their existence is certainly an obstacle to inversion of the field map without extra measurements or hypotheses (*e.g.*, measuring from below or assuming unidirectionality).

In the course of the doctoral work by D. Ponomarev, the study of the 2D spectral decomposition of the truncated Poisson operator has been undertaken. It is a simplified version of the relation between the magnetization and the magnetic potential. We considered several formulations in terms of singular integral equations and matrix Riemann-Hilbert problems, and focused on finding closed form solutions for various approximations of the Poisson operator in terms of a the ratio between the distance h to the measurement plane and the sample support size.

Lately, Apics became a partner of the ANR project MagLune, dealing with Lunar magnetism, a in collaboration with the Geophysics and Planetology Department of Cerege, CNRS, Aix-en-Provence, see section 8.2.2 . The research is just starting, and will focus on computing net moments of lunar rock samples collected by NASA.

6.2. Boundary value problems

Participants: Laurent Baratchart, Sylvain Chevillard, Juliette Leblond, Dmitry Ponomarev.

Collaboration with Laurent Bourgeois (ENSTA ParisTech, Lab. Poems), Elodie Pozzi (Univ. Bordeaux, IMB), Emmanuel Russ (Univ. Grenoble, IJF).

Generalized Hardy classes

As we mentioned in section 4.4 , 2-D diffusion equations of the form $\mathrm{div}(\sigma\nabla u) = 0$ with real non-negative valued conductivity σ can be viewed as compatibility conditions for the so-called conjugate Beltrami equation: $\overline{\partial} f = \nu \overline{\partial f}$ with $\nu = (1-\sigma)/(1+\sigma)$ [4]. Thus, the conjugate Beltrami equation is a means to replace the initial second order diffusion equation by a first order system of two real equations, merged into a single complex one. Hardy spaces under study here are those of this conjugate Beltrami equation: they are comprised of solutions to that equation in the considered domain whose L^p means over curves tending to the boundary

[0] http://www-sop.inria.fr/apics/IMPINGE/Documents/Report_Permiakova_Olga.pdf

of the domain remain bounded. They will for example replace holomorphic Hardy spaces in problem (P) when dealing with non-constant (isotropic) conductivity. Their traces merely lie in L^p $(1 < p < \infty)$, which is suitable for identification from point-wise measurements, and turn out to be dense on strict subsets of the boundary. This allows one to state Cauchy problems as bounded extremal issues in L^p classes of generalized analytic functions, in a manner which is reminiscent of what we discussed for analytic functions in section 3.3.1 .

The study of such Hardy spaces for Lipschitz σ was reduced in [4] to that of spaces of pseudo-holomorphic functions with bounded coefficients, which were apparently first considered on the disk by S. Klimentov. Solutions factorize as $e^s F$, where F is a holomorphic Hardy function while s is in the Sobolev space $W^{1,r}$ for all $r < \infty$ (Bers factorization), and the analog to the M. Riesz theorem holds which amounts to solvability of the Dirichlet problem with L^p boundary data. The case of finitely connected domains was carried out in [14].

This year, we addressed in [25] the uniqueness issue for the classical Robin inverse problem on a Lipschitz-smooth domain $\Omega \subset \mathbb{R}^n$, with L^∞ Robin coefficient, L^2 Neumann data and isotropic conductivity of class $W^{1,r}(\Omega)$, $r > n$. The Robin inverse problem consists in recovering the ratio of the normal derivative and the solution (the so-called Robin coefficient) on a subset of the boundary, knowing them on the complementary subset. We showed that uniqueness of the Robin coefficient on a subset of the boundary, given Cauchy data on the complementary subset, does hold when $n = 2$ whenever the boundary subsets are of positive Lebesgue measure. We also showed that this no longer holds in higher dimension, and we gave counterexamples when $n = 3$. The subsets in these counterexamples look very bad, and it is natural to ask whether uniqueness prevails if they have interior points. This raises an interesting open issue on harmonic gradients, namely: can a nonzero harmonic function vanish together with its normal derivative on a subset of the boundary of positive measure, and still the Robin coefficient is bounded in a neighborhood of that set? This question is worth investigating

Best constrained analytic approximation

Several questions about the behavior of solutions to the bounded extremal problem (P) in section 3.3.1 , and of some generalizations thereof, are still under study by Apics.. We considered additional interpolation constraints on the disk in problem (P), and derived new stability estimates for the solution [24]. An article is being written on the subject. Ongoing work is geared towards applications of [24]. New insight leads us to relate these results to overdetermined boundary value problems for 2D Laplace equations on irregular boundaries. This has applications in set-ups where measurements are obtained from oddly distributed sensors. Treating some of the measurements as pointwise interpolation constraints seems a reasonable strategy in comparison with interpolation of the data along a geometrically complicated boundary. Such interpolation constraints arise naturally in inverse boundary problems like plasma shaping, when some of the measurements are performed inside the chamber of the tokamak, see section 4.4 .

6.3. Matching problems and their applications - De-embedding of filters in multiplexers

Participants: Laurent Baratchart, Martine Olivi, Sanda Lefteriu, David Martinez Martinez, Fabien Seyfert.

This work has been done in collaboration with Stéphane Bila (Xlim, Limoges, France), Hussein Ezzedin (Xlim, Limoges, France), Damien Pacaud (Thales Alenia Space, Toulouse, France), Giuseppe Macchiarella (Politecnico di Milano, Milan, Italy), and Matteo Oldoni (Siae Microelettronica, Milan, Italy).

6.3.1. Matching problems and their applications

Filter synthesis is usually performed under the hypothesis that both ports of the filter are loaded on a constant resistive load (usually 50 Ohm). In complex systems, filters are however cascaded with other devices, and end up being loaded, at least at one port, on a non purely resistive frequency varying load. This is for example the case when synthesizing a multiplexer: each filter is here loaded at one of its ports on a common junction. Thus, the load is by construction non constant with the frequency, and not purely resistive either. Likewise, in an emitter-receiver, the antenna is followed by a filter. Whereas the antenna can usually be regarded as a resistive load at some frequencies, this is far from being true on the whole working band. A mismatch between the antenna and the filter, however, causes irremediable power losses, both in emission and transmission. Our goal is therefore to develop a filter synthesis method that allows to match varying loads on specific frequency bands.

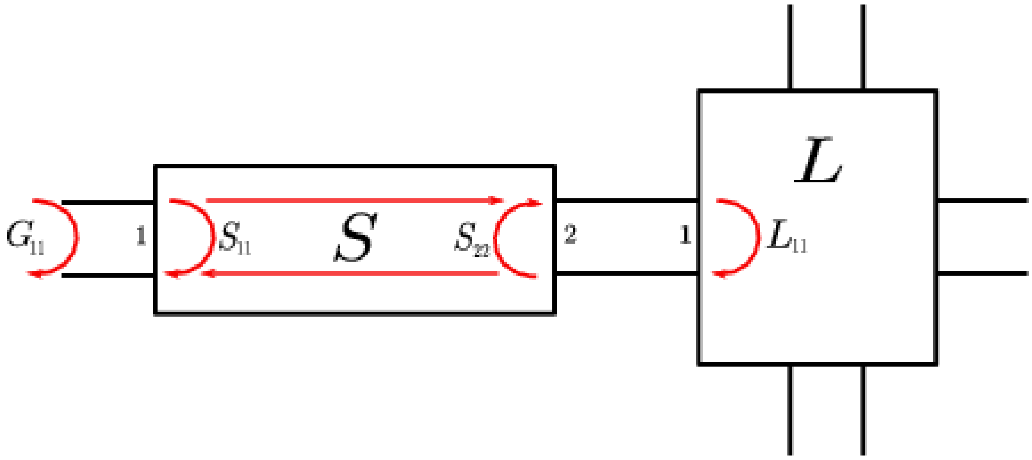


Figure 6. Filter plugged on a system with reflexion coefficient L_{11}

Figure 6 shows a filter with scattering matrix S, plugged at its right port on a frequency varying load with reflexion parameter $L_{1,1}$. If the filter is lossless, simple algebraic manipulations show that on the frequency axis the reflexion parameter satisfies:

$$|G_{1,1}| = \left| \frac{S_{2,2} - \overline{L_{1,1}}}{1 - S_{2,2}L_{1,1}} \right|.$$

The matching problem of minimizing $|G_{1,1}|$ amounts therefore to minimize the pseudo-hyperbolic distance between the filter's reflexion parameter $S_{2,2}$ and the load's reflexion $L_{1,1}$, on a given frequency band. For a broad class of filters, namely those that can be modeled by a circuit of n coupled resonators, the scattering matrix S is a rational function of McMillan degree n in the frequency. The matching problem appears therefore as a rational approximation problem in hyperbolic metric. When n is fixed, the latter is non-convex and led us to seek methods to derive good initial guesses for classical descent algorithms. To this effect, if $S_{2,2} = p/q$ we considered the following interpolation problem: given n frequency points $w_1 \cdots w_n$ and a transmission polynomial r, to find a unitary polynomial p of degree n such that:

$$j = 1..n, \qquad \frac{p}{q}(w_j) = \overline{L_{1,1}(w_j)}$$

where q is the unique monic Hurwitz polynomial of degree n satisfying the Feldtkeller equation

$$qq^* = pp^* + rr^*,$$

which accounts for the losslessness of the filter. This problem can be seen as an extended Nevanlinna-Pick interpolation problem, that was considered in [67] when the interpolation points w_j lie in the *open* left half-plane. The method in the last reference does not extend to imaginary interpolation point and we used rather different, differential-topological techniques to prove that this problem has a unique solution, which can be computed by continuation. In the setting of multiplexer synthesis, where this result must e applied recursively to each filter, we showed the existence of a fixed point for the tuning procedure, based on Brouwer's fixed point theorem. These results were presented at the MTNS [18], at the plenary of session of Ernsi workshop 2014, and they lie at the heart of the ANR Cocoram on co-integration of filters and antennas (8.2.1). Implementation of the continuation algorithm has been done under contract with CNES and yields encouraging results. Generalizations of the interpolation problem where the monic condition is relaxed are under study in the framework of co-integration of filters and antennas.

6.3.2. De-embedding of multiplexers

This work is pursued in collaboration with Thales Alenia Space, Siae Microelettronica, Xlim and under contract with CNES-Toulouse (see section 7.1).

Let S be the scattering parameters of a multiplexer composed of a N-port junction with response T and $N-1$ filters with responses $F_1, \cdots F_{N-1}$, as plotted on Figure 7 . The de-embedding problem is to recover the F_k and it can be stated under various hypotheses. Last year we studied this problem when S and T are known [79] but no special structure for the F_k is assumed. It was shown that for generic T and for $N > 3$, the de-embedding problem has a unique solution. In practice, however, the junction's response is far from being generic, as it is usually obtained *via* assembly of T-junctions. This makes the problem extremely sensitive to measurement noise. It was also noticed that in practical applications, scattering measurements of the junction are hardly available.

It is therefore natural to consider the following de-embedding problem. Given S, and under the assumption that

- the F_k are rational of known McMillan degree,
- the coupling geometry of their circuital realization is known,

what can be said about the filter's response? Note that the above assumptions do not bear on the junction. Nevertheless, we showed that the filter's responses are identifiable up to a constant matrix chained at their nearest port to the junction [73]. It was proved also that the uncertainty induced by the chain matrix bears only on the resonant frequency of the last cavity of each filter, as well as on their output coupling. Most of the filters' parameters can therefore be recovered in principle. The approach is constructive and relies on rational approximation to certain scattering parameters, as well as on some extraction procedure similar to Darlington's synthesis. Software development is under way and experimental studies have started on data provided to us by Thales Alenia Space and by Siae Microelettronica. A mid-term objective is to extend Presto-HF (see Section 5.3) so as to handle de-embedding problems for multiplexers and more generally for multi-ports.

6.4. Stability of amplifiers

Participants: Laurent Baratchart, Sylvain Chevillard, Martine Olivi, Fabien Seyfert.

This work is performed under contract with CNES-Toulouse and the University of Bilbao. The goal is to help designing amplifiers, in particular to detect instability at an early stage of the design.

Currently, electrical engineers from the University of Bilbao, under contract with CNES (the French Space Agency), use heuristics to detect instability before an amplifying circuit is physically built. Our goal is to set up a rigorously founded algorithm, based on properties of transfer functions of such amplifiers, which belong to particular classes of analytic functions.

In non-degenerate cases, non-linear electrical components can be replaced by their first order approximation when studying stability in the small signal regime. Using this approximation, diodes appear as negative resistors and transistors as current sources controlled by the voltage at certain nodes of the circuit.

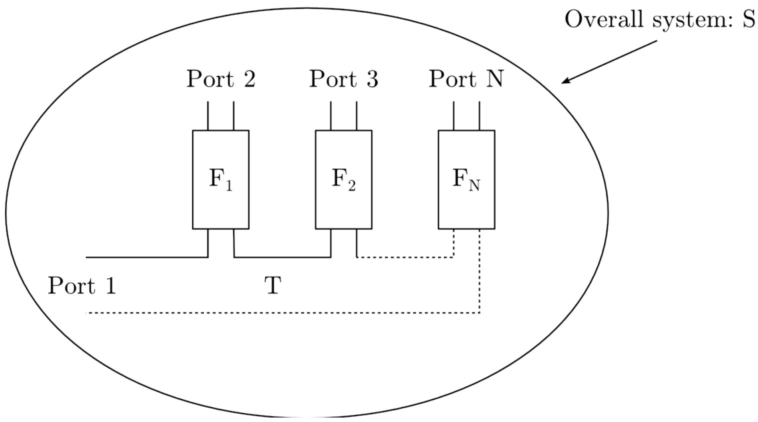


Figure 7. Multiplexer made of a junction T and filtering devices $F_1, F_2 \cdots F_N$

Over the last three years, we studied several features of transfer functions of amplifying electronic circuits:

- We characterized the class of transfer functions which can be realized with ideal components linearized active components, together with standard passive components (resistors, inductors, capacitors and transmission lines). It is exactly the field of rational functions in the complex variable and in the hyperbolic cosines and identity-times-hyperbolic sines of polynomials of degree 2 with real negative roots.
- We introduced a realistic notion of stability, by terming stable a circuit whose transfer function belongs to H^∞, as long a sufficiently high resistor is added in parallel to that circuit.
- We constructed unstable circuits having no pole in the right half-plane, which came as a surprise to our partners.
- In order to circumvent these pathological examples, we introduced a realistic hypothesis that there are small inductive and capacitive effects to active components. Our main result is that a realistic circuit without poles on the imaginary axis is unstable if and only if it has poles in the right half-plane. Moreover, there can only be finitely many of them.

This year, we were led to modify our definition of stability, taking a hint from scattering theory. We say that a transfer function Z is stable whenever $(R-Z)/(R+Z)$ belongs to H^∞ with uniformly bounded H^∞-norm for all R large enough. Equivalently, this means that the circuit can amplify signals but not require an unbounded amount of energy from the primary power circuit. This new definition is really about energy, hence is more natural. Also, it allows us a unified characterization in the corner case where instabilities are located on the imaginary axis. We obtained this way a nice characterization: Z is stable if and only if it has no pole in the open right half plane, while each pole it may have on the imaginary axis is simple and has a residue with strictly positive real part. We published a research report [23] and an article is being written to report on our results.

6.5. Approximation

Participant: Laurent Baratchart.

6.5.1. Orthogonal Polynomials

This is joint work with Nikos Stylianopoulos (Univ. of Cyprus).

We study the asymptotic behavior of weighted orthogonal polynomials on a bounded simply connected plane domain Ω. The n-th orthogonal polynomial P_n has degree n, positive leading coefficient, and satisfies

$$\int_\Omega P_n \overline{P}_k w \, dm = \delta_{n,k}$$

where w is an integrable positive weight and $\delta_{n,k}$ is the Kronecker symbol. When Ω is smooth while w is Hölder-continuous and non-vanishing, it is known that

$$P_n(z) = \left(\frac{n+1}{\pi}\right)^{1/2} \Phi^n \frac{\Phi'}{S_w(z)} \{1 + o(1)\},$$

locally uniformly outside the convex hull of $\overline{\Omega}$, where Φ is the conformal map from the complement of Ω onto the complement of the unit disk and S_w is the so-called Szegö function of the trace of w on the boundary $\partial\Omega$ [81]. If we compare it with classical exterior Szegő asymptotics, the formula asserts that P_n behaves asymptotically like the n-th orthogonal polynomial with respect to a weight supported $on\partial\Omega$ (the trace of w), up to a factor $\sqrt{(n+1)/\pi}$.

When Ω is the unit disk, we proved this result under unprecedented weak assumptions on w, namely $w(re^{i\theta})$ should converge in $L^p(T)$ as $r \to 1$ for some $p > 1$ and its $\log^-$ should be bounded in the real Hardy space H^1. An article is being written on these findings and the case of a smooth domain Ω, more general than a disk, is under examination.

6.5.2. Meromorphic approximation

This is joint work with Maxim Yattselev (Purdue Univ. at Indianapolis, USA).

We proved in [6] that the normalized counting measure of poles of best H^2 approximants of degree n to a function analytically continuable, except over finitely many branchpoints lying outside the unit disk, converges to the Green equilibrium distribution of the compact set of minimal Green capacity outside of which the function is single valued (the normalized counting measure is the probability measure with equal mass at each pole). This result warrants source recovery techniques used in section 6.1.1 . Here we consider the corresponding problem for best uniform meromorphic approximants on the unit circle (so-called AAK approximants after Adamjan, Arov and Krein), in the case where the function may have poles and essential singularities. This year, we established a similar result when the function has finitely many essential singularities. The general case is still pending.

ARAMIS Project-Team

6. New Results

6.1. Highlights of the Year

ARAMIS has contributed to the special issue on "Complex network theory and the brain" in the prestigious journal of Philosopical Transactions of the Royal Society, Series B. This work was featured by the ICM (http://icm-institute.org/en/news/complex-network-theory-and-the-brain?lang=en) and Inria (http://www.inria.fr/en/centre/paris-rocquencourt/news/complex-network-theory-and-the-brain).

6.2. Detection of volume loss in hippocampal layers in Alzheimer's disease using 7 T MRI

Participants: Claire Boutet, Marie Chupin, Stéphane Lehéricy, Linda Marrakchi-Kacem, Stéphane Epelbaum, Cyril Poupon, Christopher Wiggins, Alexandre Vignaud, Dominique Hasboun, Bénédicte Desfontaines, Olivier Hanon, Bruno Dubois, Marie Sarazin, Lucie Hertz-Pannier, Olivier Colliot [Correspondant].

In Alzheimer's disease (AD), the hippocampus is an early site of tau pathology and neurodegeneration. Histological studies have shown that lesions are not uniformly distributed within the hippocampus. Moreover, alterations of different hippocampal layers may reflect distinct pathological processes. 7 T MRI dramatically improves the visualization of hippocampal subregions and layers. In this study, we aimed to assess whether 7 T MRI can detect volumetric changes in hippocampal layers in vivo in patients with AD. We studied four AD patients and seven control subjects. MR images were acquired using a whole-body 7 T scanner with an eight channel transmit-receive coil. Hippocampal subregions were manually segmented from coronal T2*-weighted gradient echo images with $0.3 \times 0.3 \times 1.2$ mm3 resolution using a protocol that distinguishes between layers richer or poorer in neuronal bodies (Figure 1). Five subregions were segmented in the region of the hippocampal body: alveus, strata radiatum, lacunosum and moleculare (SRLM) of the cornu Ammonis (CA), hilum, stratum pyramidale of CA and stratum pyramidale of the subiculum. We found strong bilateral reductions in the SRLM of the cornu Ammonis and in the stratum pyramidale of the subiculum ($p < 0.05$), with average cross-sectional area reductions ranging from -29% to -49%. These results show that it is possible to detect volume loss in distinct hippocampal layers using segmentation of 7 T MRI. 7 T MRI-based segmentation is a promising tool for AD research.

More details in [3].

6.3. White matter lesions in patients with frontotemporal lobar degeneration due to progranulin mutations

Participants: Paola Caroppo, Isabelle Le Ber, Agnès Camuzat, Fabienne Clot, Lionel Naccache, Foudil Lamari, Anne Bertrand, Serge Belliard, Olivier Colliot [Correspondant], Alexis Brice.

Mutations in the progranulin (GRN) gene are responsible for 20% of familial cases of frontotemporal dementias. All cause haploinsufficiency of progranulin, a protein involved in inflammation, tissue repair, and cancer. Carriers of the GRN mutation are characterized by a variable degree of asymmetric brain atrophy, predominantly in the frontal, temporal, and parietal lobes. We described four GRN mutation carriers with remarkable widespread white matter lesions (WML) associated with lobar atrophy shown on magnetic resonance imaging. The WML were predominantly in the frontal and parietal lobes and were mostly confluent, affecting the periventricular subcortical white matter and U-fibers. In all patients, common vascular, metabolic, inflammatory, dysimmune, and mitochondrial disorders were excluded and none had severe vascular risk factors. Our data suggest that white matter involvement may be linked to progranulin pathological processes in a subset of GRN mutation carriers. The plasma progranulin measurement, which is predictive of GRN mutations, and GRN sequencing should thus be included in investigations of patients with frontotemporal lobar degenerations who show unusual white matter hyperintensities and atrophy on magnetic resonance imaging.

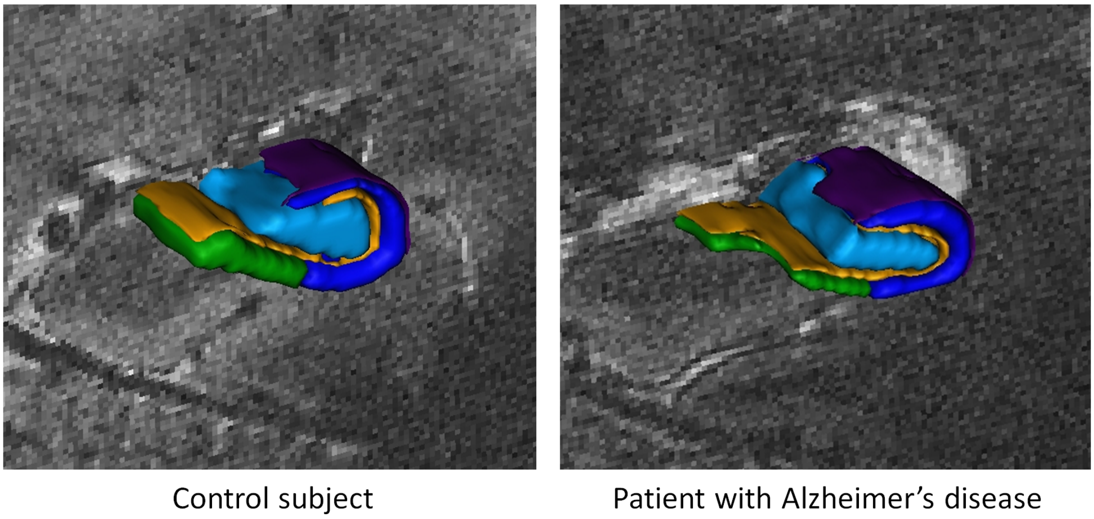


Figure 1. Segmentation of hippocampal layers using in vivo 7 Tesla MRI. were performed on the second echo image. Left panel: control subject. Right panel: patient with Alzheimer’s disease. Purple, alveus; dark blue, stratum pyramidale of CA1-3; yellow, strata radiatum,lacunosum and moleculare of CA1-3, strata lacunosum and moleculare of the subiculum and stratum moleculare of gyrus dentatus; cyan, stratum pyramidale of CA4 and stratum granulosum and polymorphic layer of gyrus dentatus; green, stratum pyramidale of the subiculum.

More details in [4].

6.4. Template-based morphometry using diffeomorphic iterative centroids

Participants: Claire Cury [Correspondant], Joan Glaunès, Marie Chupin, Olivier Colliot.

A common approach for the analysis of anatomical variability relies on the estimation of a representative template of the population, followed by the study of this population based on the parameters of the deformations going from the template to the population. The Large Deformation Diffeomorphic Metric Mapping framework is widely used for shape analysis of anatomical structures, but computing a template with such framework is computationally expensive. We proposed a fast approach for template-based analysis of anatomical variability. The template is estimated using an iterative approach which quickly provides a centroid of the population. Statistical analysis is then performed using principal component analysis on the initial momenta that define the deformations between the centroid and each subject of the population. This approach was applied to the analysis of hippocampal shape on 80 patients with Alzheimer's Disease and 138 controls from the ADNI database.

More details in [22] and [36].

6.5. Structural connectivity differences in left and right temporal lobe epilepsy

Participants: Pierre Besson, Vera Dinkelacker [Correspondant], Romain Valabrègue, Lionel Thivard, Xavier Leclerc, Michel Baulac, Daniela Sammler, Olivier Colliot, Stéphane Lehéricy, Séverine Samson, Sophie Dupont.

Our knowledge on temporal lobe epilepsy (TLE) with hippocampal sclerosis has evolved towards the view that this syndrome affects widespread brain networks. Diffusion weighted imaging studies have shown alterations of large white matter tracts, most notably in left temporal lobe epilepsy, but the degree of altered connections between cortical and subcortical structures remains to be clarified. We performed a whole brain connectome analysis in 39 patients with refractory temporal lobe epilepsy and unilateral hippocampal sclerosis (20 right and 19 left) and 28 healthy subjects. We performed whole-brain probabilistic fiber tracking using MRtrix and segmented 164 cortical and subcortical structures with Freesurfer. Individual structural connectivity graphs based on these 164 nodes were computed by mapping the mean fractional anisotropy (FA) onto each tract. Connectomes were then compared using two complementary methods: permutation tests for pair-wise connections and Network Based Statistics to probe for differences in large network components. Comparison of pair-wise connections revealed a marked reduction of connectivity between left TLE patients and controls, which was strongly lateralized to the ipsilateral temporal lobe. Specifically, infero-lateral cortex and temporal pole were strongly affected, and so was the perisylvian cortex. In contrast, for right TLE, focal connectivity loss was much less pronounced and restricted to bilateral limbic structures and right temporal cortex. Analysis of large network components revealed furthermore that both left and right hippocampal sclerosis affected diffuse global and interhemispheric connectivity. Thus, left temporal lobe epilepsy was associated with a much more pronounced pattern of reduced FA, that included major landmarks of perisylvian language circuitry. These distinct patterns of connectivity associated with unilateral hippocampal sclerosis show how a focal pathology influences global network architecture, and how left or right-sided lesions may have differential and specific impacts on cerebral connectivity.

More details in [2].

6.6. Morphometry of anatomical shape complexes with dense deformations and sparse parameters

Participants: Stanley Durrleman [Correspondant], Marcel Prastawa, Nicolas Charon, Julie Korenberg, Sarang Joshi, Guido Gerig, Alain Trouvé.

We propose a generic method for the statistical analysis of collections of anatomical shape complexes, namely sets of surfaces that were previously segmented and labeled in a group of subjects. The method estimates an anatomical model, the template complex, that is representative of the population under study. Its shape reflects anatomical invariants within the dataset. In addition, the method automatically places control points near the most variable parts of the template complex. Vectors attached to these points are parameters of deformations of the ambient 3D space. These deformations warp the template to each subject's complex in a way that preserves the organization of the anatomical structures. Multivariate statistical analysis is applied to these deformation parameters to test for group differences. Results of the statistical analysis are then expressed in terms of deformation patterns of the template complex, and can be visualized and interpreted. The user needs only to specify the topology of the template complex and the number of control points. The method then automatically estimates the shape of the template complex, the optimal position of control points and deformation parameters. The proposed approach is completely generic with respect to any type of application and well adapted to efficient use in clinical studies, in that it does not require point correspondence across surfaces and is robust to mesh imperfections such as holes, spikes, inconsistent orientation or irregular meshing.

The approach is illustrated with a neuroimaging study of Down syndrome (DS). Results demonstrate that the complex of deep brain structures shows a statistically significant shape difference between control and DS subjects. The deformation-based modeling is able to classify subjects with very high specificity and sensitivity, thus showing important generalization capability even given a low sample size. We show that results remain significant even if the number of control points, and hence the dimension of variables in the statistical model, are drastically reduced. The analysis may even suggest that parsimonious models have an increased statistical performance.

The method has been implemented in the software Deformetrica, which is publicly available at www.deformetrica.org

More details in [14].

6.7. Iconic-Geometric Nonlinear Registration of a Basal Ganglia Atlas for Deep Brain Stimulation Planning

Participants: Ana Fouquier, Stanley Durrleman, Jérôme Yelnik, Sara Fernandez-Vidal, Eric Bardinet.

We evaluated a nonlinear registration method for warping a 3D histological atlas of the basal ganglia into patient data for deep brain stimulation (DBS) planning. The power of the method is the possibility to combine iconic registration with geometric constraints under a unified diffeomorphic framework. This combination aims to ensure robust and accurate atlas-to-patient warping and anatomy-preserving deformations of stimulation target nuclei. A comparison of the method with a state-of-the-art diffeomorphic registration algorithm reveals how each approach deforms low-contrasted image regions where DBS target nuclei often lie. The technique is applied to T1-weighted magnetic resonance images from a cohort of Parkinsonian subjects, including subjects with standard-size and large ventricles. Results illustrate the effects of iconic or geometric registration alone, as well as how both constraints can be integrated in order to contribute for registration precision enhancement. See Fig. 2 .

More details in [25].

6.8. Evaluation of morphometric descriptors of deep brain structures for the automatic classification of patients with Alzheimer's disease, mild cognitive impairment and elderly controls

Participants: Alexandre Routier [correspondant], Pietro Gori, Ana Fouquier, Sophie Lecomte, Olivier Colliot, Stanley Durrleman.

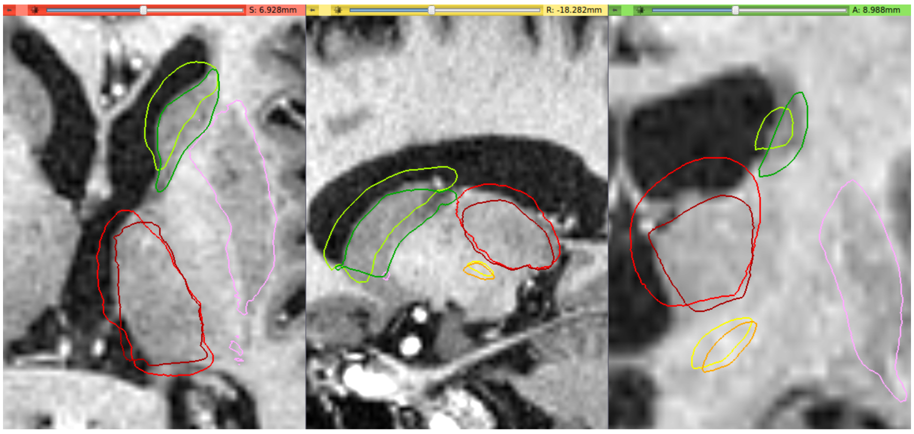


Figure 2. Superimposition of deformed meshes of the histological atlas with a patient pre-operative MRI. Meshes in bright colors result from a block-matching algorithm based on image intensity. Meshes in dark colors result from our iconic-geometric approach with non-linear deformation. We observe a better alignment of the structures, as well as a realistic deformation of the sub-thalamic nucleus (in yellow/orange), which is not visible in the image and therefore has not been taken into account for estimating the optimal deformation. This nucleus is the stimulation target for patients with Parkinson's disease.

We participated in the Computer-Aided Diagnosis of Dementia based on structural MRI data (http://caddementia.grand-challenge.org/). Our approach was to select shapes of 12 brain structures: the caudate nucleus, putamen, pallidum, thalamus, hippocampus and amygdala of each hemisphere. The structure segmentation was based on a FreeSurfer segmentation and the marching-cubes algorithm was used to get 3D triangular meshes. Using our software Deformetrica, anatomical models (mean shape and typical variations) of these brain structures were built for patients with Alzheimer's disease (AD), Mild Cognitive Impairments (MCI) and cognitively normal controls (CN) based on the data of 509 ADNI subjects. The models for AD, MCI and CN were registered to the test subjects by maximizing the likelihood of the test image to be derived from each model. The final classification was made by thresholding this criterion taking into account the covariance of the deformation parameters. The thresholds were either optimized on the ADNI data or on the provided training data. The method was fully automatic and the computation time was 4 days for training the anatomical models plus 11 hours per subject for registration and classification. For the 30 training subjects, the algorithm had accuracies of 73% (if optimized on training data) and 50% (if optimized on ADNI data). On the test set of 354 images, our method yields an accuracy of 49.2% (43.5 - 54.2), true positive fraction of 94.6% (89.8 - 97.7) for CN, 11.5% (6.2 - 17.7) for MCI and 36.9% (27.4 - 46.5) for AD.

Our participation to this challenge was the opportunity to test our software Deformetrica for classification tasks. It ran on more than 800 images, thus showing its ability to deal with large data sets.

More details in [27].

6.9. A Prototype Representation to Approximate White Matter Bundles with Weighted Currents

Participants: Pietro Gori [correspondant], Olivier Colliot, Linda Marrakchi-Kacem, Fabrizio de Vico Fallani, Mario Chavez, Sophie Lecomte, Cyril Poupon, Andreas Hartmann, Nicholas Ayache, Stanley Durrleman.

Quantitative and qualitative analysis of white matter fibers resulting from tractography algorithms is made difficult by their huge number. To this end, we propose an approximation scheme which gives as result a more concise but at the same time exhaustive representation of a fiber bundle. It is based on a novel computational model for fibers, called weighted currents, characterized by a metric that considers both the pathway and the anatomical locations of the endpoints of the fibers. Similarity has therefore a twofold connotation: geometrical and related to the connectivity. The core idea is to use this metric for approximating a fiber bundle with a set of weighted prototypes, chosen among the fibers, which represent ensembles of similar fibers. The weights are related to the number of fibers represented by the prototypes. The algorithm is divided into two steps. First, the main modes of the fiber bundle are detected using a modularity based clustering algorithm. Second, a prototype fiber selection process is carried on in each cluster separately. This permits to explain the main patterns of the fiber bundle in a fast and accurate way. See Fig. 3

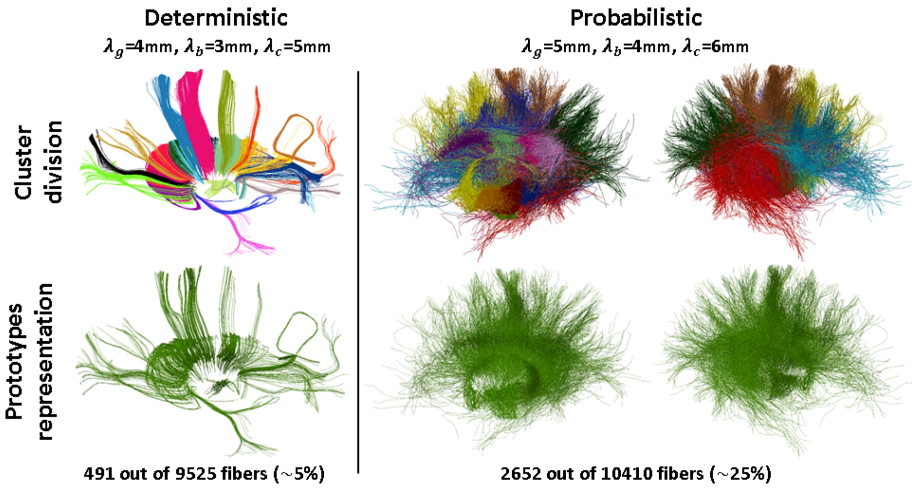


Figure 3. Illustration of our method to cluster fibers and approximate clusters based on a weighted currents metric, which measures differences in the locations of fibers extremities and the geometry of their pathway. 2 examples are shown using fibers from a deterministic tractography (left) and probabilistic tractography (right). Clustering (top row) and approximation of fibers within each cluster (bottom row) are shown.

More details in [24].

6.10. Non-parametric resampling of random walks for spectral network clustering

Participants: Fabrizio de Vico Fallani [correspondant], Vincenzo Nicosia, Vito Latora, Mario Chavez.

Parametric resampling schemes have been recently introduced in complex network analysis with the aim of assessing the statistical significance of graph clustering and the robustness of community partitions. We proposed a method to replicate structural features of complex networks based on the non-parametric resampling of the transition matrix associated with an unbiased random walk on the graph. We tested this bootstrapping technique on synthetic and real-world modular networks and we showed that the ensemble of replicates obtained through resampling can be used to improve the performance of standard spectral algorithms for community detection.

More details in [10].

6.11. Graph analysis of functional brain networks: practical issues in translational neurosciences

Participants: Fabrizio de Vico Fallani [correspondant], Sophie Achard, Jonas Richiardi, Mario Chavez.

The brain can be regarded as a network: a connected system where nodes, or units, represent different specialized regions and links, or connections, represent communication pathways. From a functional perspective, communication is coded by temporal dependence between the activities of different brain areas. In the last decade, the abstract representation of the brain as a graph has allowed to visualize functional brain networks and describe their non-trivial topological properties in a compact and objective way. Nowadays, the use of graph analysis in translational neuroscience has become essential to quantify brain dysfunctions in terms of aberrant reconfiguration of functional brain networks. Despite its evident impact, graph analysis of functional brain networks is not a simple toolbox that can be blindly applied to brain signals. On the one hand, it requires the know-how of all the methodological steps of the pipeline that manipulate the input brain signals and extract the functional network properties. On the other hand, knowledge of the neural phenomenon under study is required to perform physiologically relevant analysis. The aim of our work is to provide practical indications to make sense of brain network analysis and contrast counterproductive attitudes.

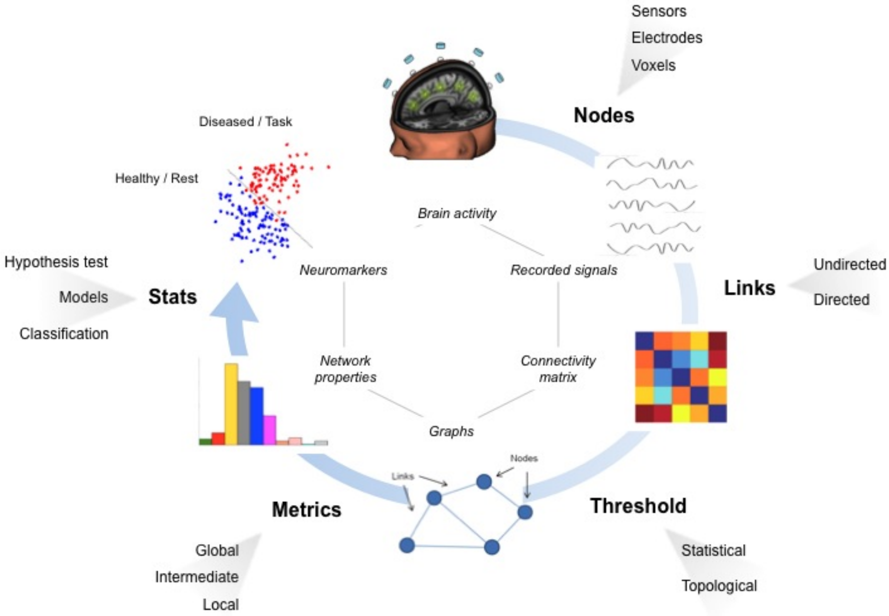


Figure 4. Processing pipeline for functional brain connectivity modeling and analysis. Nodes correspond to specific brain sites according to the used neuroimaging technique. Links are estimated by measuring the functional connectivity (FC) between the activity of brain nodes; this information is contained in a connectivity matrix. By means of filtering procedures, based on thresholds, only the most important links constitute the brain graph. The topology of the brain graph is quantified by different graph metrics (or indices) that can be represented as numbers (e.g. the colored bars). These graph indices can be input to statistical analysis to look for significant differences between populations/conditions (e.g. red points correspond to brain graph indices of diseased patients or tasks, blue points stand for healthy subjects or resting states).

More details in [11].

6.12. Hierarchy of neural organisation in the zebra fish spinal cord: causality analysis of in-vivo calcium imaging data

Participants: Fabrizio de Vico Fallani [correspondant], Martina Corazzol, Jnena Sternberg, Kevin Fidelin, Claire Wyart, Mario Chavez.

The recent development of genetically encoded calcium indicators enables monitoring in vivo the activity of neuronal populations. Most analysis of these calcium transients relies on linear regression analysis based on the sensory stimulus applied or the behavior observed. To estimate the basic properties of the functional neural circuitry, we propose a network-based approach based on calcium imaging recorded at single cell resolution. Differently from previous analysis based on cross-correlation, we used Granger causality estimates to infer activity propagation between the activities of different neurons. The resulting functional networks were then modeled as directed graphs and characterized in terms of connectivity and node centralities. We applied our approach to calcium transients recorded at low frequency (4 Hz) in ventral neurons of the zebrafish spinal cord at the embryonic stage when spontaneous coiling of the tail occurs. Our analysis on population calcium imaging data revealed a strong ipsilateral connectivity and a characteristic hierarchical organization of the network hubs that supported established propagation of activity from rostral to caudal spinal cord. Our method could be used for detecting functional defects in neuronal circuitry during development and pathological conditions.

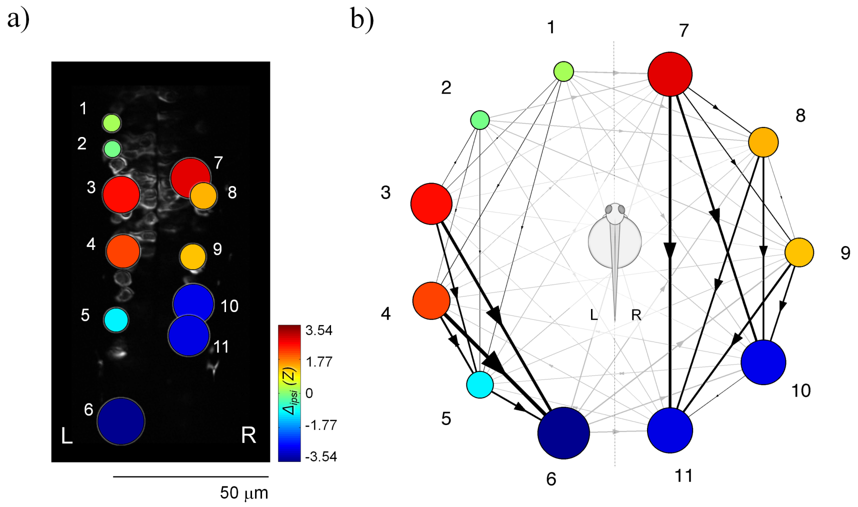


Figure 5. Rostro-caudal distribution of the nodal delta centrality in the representative zebrafish embryo. Panel a) The normalized ipsi value is represented for each node (motoneuron) as a colored circle superimposed on the field of view. The larger the circle, the more central is the node in terms of its tendency to act as a transmitter (red color, positive value) or receiver (blue color, negative value) hub of information flow. Panel b) The same normalized ipsi centrality values are here represented within the neuronal GC network. Statistically significant GC influences are illustrated as directed arrows. The thicker the arrow the stronger the GC value is. Inter-hemicord directed links are illustrated in gray color for the sake of simplicity.

More details in [9].

6.13. 2D harmonic filtering of MR phase images in multicenter clinical setting: towards a magnetic signature of cerebral microbleeds

Participants: Takoua Kaaouana [correspondant], Ludovic de Rochefort, Thomas Samaille, Nathalie Thiery, Carole Dufouil, Christine Delmaire, Didier Dormont, Marie Chupin.

Cerebral microbleeds (CMBs) have emerged as a new imaging marker of small vessel disease. Composed of hemosiderin, CMBs are paramagnetic and can be detected with MRI sequences sensitive to magnetic susceptibility (typically, gradient recalled echo T2* weighted images). Nevertheless, their identification remains challenging on T2* magnitude images because of confounding structures and lesions. In this context, T2* phase image may play a key role in better characterizing CMBs because of its direct relationship with local magnetic field variations due to magnetic susceptibility difference. To address this issue, susceptibility-based imaging techniques were proposed, such as Susceptibility Weighted Imaging (SWI) and Quantitative Susceptibility Mapping (QSM). But these techniques have not yet been validated for 2D clinical data in multicenter settings. Here, we introduce 2DHF, a fast 2D phase processing technique embedding both unwrapping and harmonic filtering designed for data acquired in 2D, even with slice-to-slice inconsistencies. This method results in internal field maps which reveal local field details due to magnetic inhomogeneity within the region of interest only. This technique is based on the physical properties of the induced magnetic field and should yield consistent results. A synthetic phantom was created for numerical simulations. It simulates paramagnetic and diamagnetic lesions within a "brain-like' tissue, within a background. The method was evaluated on both this synthetic phantom and multicenter 2D datasets acquired in a standardized clinical setting, and compared with two state-of-the-art methods. It proved to yield consistent results on synthetic images and to be applicable and robust on patient data. As a proof-of-concept, we finally illustrate that it is possible to find a magnetic signature of CMBs and CMCs on internal field maps generated with 2DHF on 2D clinical datasets that gives consistent results with CT-scans in a subsample of 10 subjects acquired with both modalities. See Fig. 6

More details in [16].

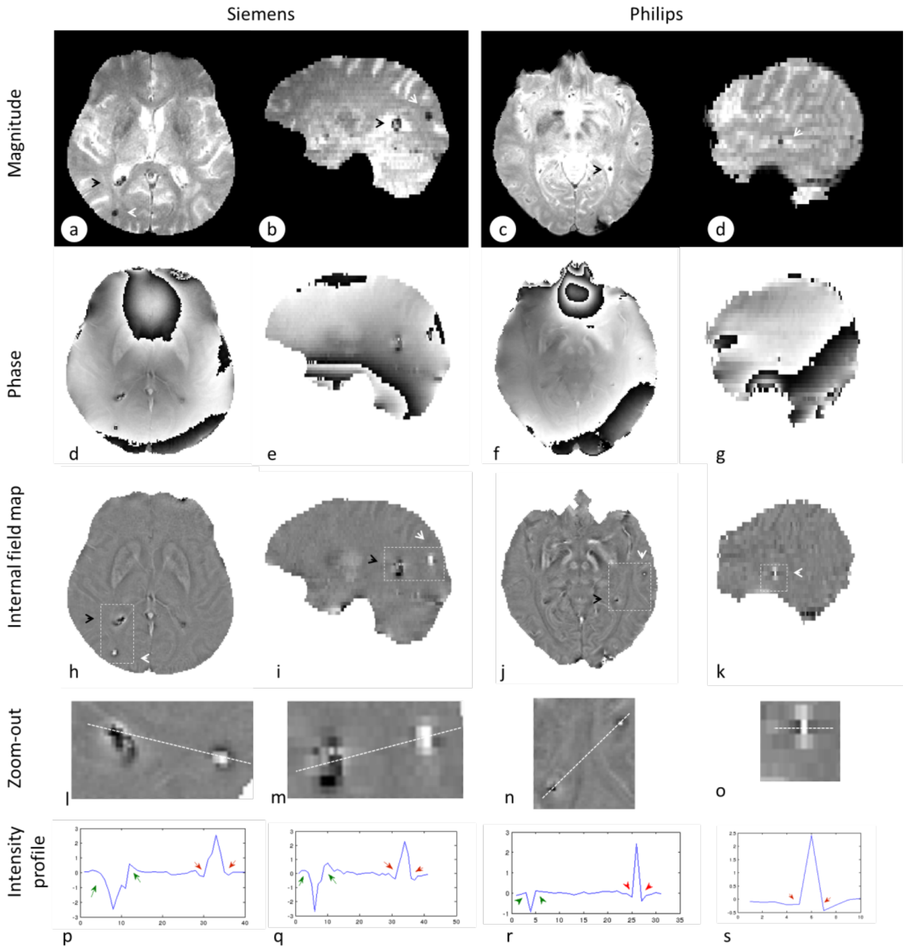


Figure 6. Siemens (left) and Philips (right) axial and sagittal views. Magnitude image (first row), native phase image (second row) and internal field map (third row). Fourth row shows a zoomed out region corresponding to the white rectangle showing CMB with a dipolar pattern (white arrow) and a physiologic calcification of the choroid plexus (black arrow). Note that panel l was rotated. A 1D intensity profile calculated through CMBs and calcification in the zoomed region is displayed in the last row. Note the intensity sign inversion for both side of CMBs (red arrow head), and the calcification (green arrow head). Double headed arrows on panels (l-o) indicate the location of the lines used to generate the intensity profiles.

ARIC Project-Team

6. New Results

6.1. Arithmetic operators

6.1.1. A table-based method to evaluate trigonometric functions

Linear (order-one) function evaluation schemes, such as bipartite and multipartite tables, are usually effective for low precision approximations. For high output precision, the lookup table size is often too large for practical use. Dong Wang and Milos Ercegovac (UC Los Angeles) and Nicolas Brisebarre and Jean-Michel Muller investigate the so-called (M, p, k) scheme that reduces the range of input argument to a very small interval so that trigonometric functions can be approximated with very small lookup tables and a few additions/subtractions. An optimized hardware architecture is proposed and implemented in both FPGA device and standard cell based technology. Experimental results show that the proposed scheme achieves more than 50% reduction in total chip area compared with the best linear approach for 24-bit evaluation [14].

6.2. Floating-Point arithmetic

6.2.1. On the computation of the reciprocal of floating point expansions using an adapted Newton-Raphson iteration

Many numerical problems require a higher computing precision than that offered by common floating point (FP) formats. One common way of extending the precision is to represent numbers in a *multiple component* format. With so-called *floating point expansions*, numbers are represented as the unevaluated sum of standard machine precision FP numbers. This format offers the simplicity of using directly available and highly optimized FP operations and is used by multiple-precisions libraries such as Bailey's 'Q'D or the analogue Graphics Processing Units tuned version, GQD. Mioara Joldes (LAAS), Jean-Michel Muller, and Valentina Popescu introduced a new algorithm for computing the reciprocal FP expansion a^{-1} of a FP expansion a. Their algorithm is based on using an adapted Newton-Raphson iteration where "truncated" operations (additions, multiplications) involving FP expansions are used. The error analysis given shows that their algorithm allows for computations of very accurate quotients. Precisely, after $i \geq 0$ iterations, the computed FP expansion $x = x_0 + \cdots + x_{2^i-1}$ satisfies the relative error bound $|\frac{x-a^{-1}}{a^{-1}}| \leq 2^{-2^i(p-3)-1}$, where $p > 4$ is the precision of the FP representation used ($p = 24$ for single precision and $p = 53$ for double precision) [19].

6.2.2. Error bounds on complex floating-point multiplication with a fused-multiply add

The accuracy analysis of complex floating-point multiplication done by Brent, Percival, and Zimmermann [*Math. Comp.*, 76:1469–1481, 2007] is extended by Peter Kornerup (Odense Univ. Denmark), Claude-Pierre Jeannerod, Nicolas Louvet, and Jean-Michel Muller [42] to the case where a fused multiply-add (FMA) operation is available. Considering floating-point arithmetic with rounding to nearest and unit roundoff u, they show that the bound $\sqrt{5}\,u$ on the normwise relative error $|\widehat{z}/z - 1|$ of a complex product z can be decreased further to $2u$ when using the FMA in the most naive way. Furthermore, they prove that the term $2u$ is asymptotically optimal not only for this naive FMA-based algorithm, but also for two other algorithms, which use the FMA operation as an efficient way of implementing rounding error compensation. Thus, although highly accurate in the componentwise sense, these two compensated algorithms bring no improvement to the normwise accuracy $2u$ already achieved using the FMA naively. Asymptotic optimality is established for each algorithm thanks to the explicit construction of floating-point inputs for which it is proven that the normwise relative error then generated satisfies $|\widehat{z}/z - 1| \to 2u$ as $u \to 0$. All these results hold for IEEE floating-point arithmetic, with radix $\beta \geq 2$, precision $p \geq 2$, and rounding to nearest; it is only assumed that underflows and overflows do not occur and, when bounding errors from below, that $\beta^{p-1} \geq 12$.

6.2.3. Refined error analysis of the Cornea-Harrison-Tang method for $ab + cd$

In their book *Scientific Computing on Itanium-based Systems*, Cornea, Harrison, and Tang introduced an accurate algorithm for evaluating expressions of the form $ab + cd$ in binary floating-point arithmetic, assuming a fused-multiply add instruction is available. They showed that if p is the precision of the floating-point format and if $u = 2^{-p}$, the relative error of the result is of order u. Jean-Michel Muller improved their proof to show that the relative error is bounded by $2u + 7u^2 + 6u^3$. Furthermore, by building an example for which the relative error is asymptotically (as $p \to \infty$ or, equivalently, as $u \to 0$) equivalent to $2u$, he proved that this error bound is asymptotically optimal [11]. Claude-Pierre Jeannerod then showed in [41] that an error bound of the form $2u + 2u^2 + O(u^3)$ in fact holds for any radix $\beta \geq 2$, with $u = \frac{1}{2}\beta^{1-p}$. He also showed that the possibility of removing the $O(u^2)$ term from this bound depends on the radix parity and the tie-breaking strategy used for rounding: if β is odd or rounding is *to nearest even* then the simpler bound $2u$ is obtained, while if β is even and rounding is *to nearest away*, then there exist floating-point inputs a, b, c, d that lead to a relative error larger than $2u + \frac{1}{\beta}u^2$.

6.2.4. On the maximum relative error when computing integer powers by iterated multiplications in floating-point arithmetic

Stef Graillat (Paris 6 University), Vincent Lefèvre and Jean-Michel Muller improved the usual relative error bound for the computation of x^n through iterated multiplications by x in binary floating-point arithmetic. The obtained error bound is only slightly better than the usual one, but it is simpler. They also discussed the more general problem of computing the product of n terms [7].

6.2.5. Improved error bounds for numerical linear algebra

When computing matrix factorizations and solving linear systems in floating-point arithmetic, classical rounding error analyses provide backward error bounds whose leading terms have the form $\gamma_n = nu/(1 - nu)$ for suitable values of n and with u the unit roundoff. With Siegfried M. Rump (Hamburg University of Technology), Claude-Pierre Jeannerod showed in [13] that for LU and Cholesky factorizations as well as for triangular system solving, γ_n can be replaced by the $O(u^2)$-free and unconditional constant nu. To get these new bounds the main ingredient is a general framework for bounding expressions of the form $|\rho - s|$, where s is the exact sum of a floating-point number and $n - 1$ real numbers, and where ρ is a real number approximating the computed sum $\widehat{s}$.

6.2.6. On relative errors of floating-point operations

Rounding error analyses of numerical algorithms are most often carried out via repeated applications of the so-called standard models of floating-point arithmetic. Given a round-to-nearest function RN and barring underflow and overflow, such models bound the relative errors $E_1(t) = |t - \mathrm{RN}(t)|/|t|$ and $E_2(t) = |t - \mathrm{RN}(t)|/|\mathrm{RN}(t)|$ by the unit roundoff u. In [34] Claude-Pierre Jeannerod and Siegfried M. Rump (Hamburg University of Technology) investigated the possibility of refining these bounds, both in the case of an arbitrary real t and in the case where t is the exact result of an arithmetic operation on some floating-point numbers. They provided explicit and attainable bounds on $E_1(t)$, which are all less than or equal to $u/(1 + u)$ and, therefore, smaller than u. For $E_2(t)$ the bound u is attainable whenever $t = x \pm y$ or $t = xy$ or, in base > 2, $t = x/y$ with x, y two floating-point numbers. However, for division in base 2 as well as for square root, smaller bounds are derived, which are also shown to be attainable. This set of sharp bounds was then applied to the rounding error analysis of various numerical algorithms: in all cases, they obtained either much shorter proofs of the best-known error bounds for such algorithms, or improvements on these bounds themselves.

6.2.7. Comparison between binary and decimal floating-point numbers

In collaboration with Christoph Lauter and Marc Mezzarobba (LIP6 laboratory, Paris), Nicolas Brisebarre and Jean-Michel Muller introduce an algorithm to compare a binary floating-point (FP) number and a decimal FP number, assuming the “binary encoding” of the decimal formats is used, and with a special emphasis on the basic interchange formats specified by the IEEE 754-2008 standard for FP arithmetic. It is a two-step algorithm: a first pass, based on the exponents only, quickly eliminates most cases, then, when the first pass does not suffice, a more accurate second pass is performed. They provide an implementation of several variants of our algorithm, and compare them [37].

6.2.8. Correctly rounded sum of floating-point numbers in GNU MPFR

Vincent Lefèvre has designed a new algorithm to compute the correctly rounded sum of several floating-point numbers, each having its own precision and the output having its own precision, as in GNU MPFR. At the same time, the `mpfr_sum` function is being reimplemented (not finished yet). While the old algorithm was just an application of Ziv's method, thus with exponential time and memory complexity in the worst case such as the sum of a huge number and a tiny number, the new algorithm does the sum by blocks (reiterations being needed only in case of cancellations), taking such holes between numbers into account.

6.3. Certified computing and computer algebra

6.3.1. Standardization of interval arithmetic

The IEEE 1788 working group is devoted to the standardization of interval arithmetic. V. Lefèvre and N. Revol are very active in this group. This year has been devoted to a ballot on the whole text of the standard [28], and to editorial work to make it compliant with IEEE rules. The final, remaining step, is the so-called "Sponsor ballot" and it should be completed in 2015.

6.3.2. Interval linear algebra on multi-core processors

For the product of matrices with interval coefficients, fast approximate algorithms have been developed by Philippe Théveny: they compute an enclosure of the exact product. These algorithms rely on the representation of intervals by their midpoints and radii. This representation allows one to use optimized routines for the multiplication of matrices with floating-point coefficients. In [4], the quality of the approximation of several algorithms is established, which accounts for roundoff errors and not only method's errors. A new algorithm is proposed, which requires even less (only 2) calls to a floating-point routine and still offers a good approximation quality, for a well specified type of input matrices. Three of the studied algorithms are implemented on a multi-core architecture. To avoid problems listed in [12] and to offer good performances, Philippe Théveny developed optimizations. The resulting implementations exhibit good performances: guaranteed results are obtained with an overhead less than 3, high numerical intensity and good scalability.

6.3.3. Numerical reproducibility

What is called *numerical reproducibility* is the problem of getting the same result when the scientific computation is run several times, either on the same machine or on different machines. In [12], the focus is on interval computations using floating-point arithmetic: Nathalie Revol and Philippe Théveny identified implementation issues that may invalidate the inclusion property, and presented several ways to preserve this inclusion property. This work has also been replaced in the larger context of numerical validation [15].

6.3.4. Faster multivariate interpolation with multiplicities

Muhammad Chowdhury (U. Western Ontario), Claude-Pierre Jeannerod, Vincent Neiger (ENS de Lyon), Éric Schost (U. Western Ontario), and Gilles Villard proposed in [38] a fast algorithm for interpolating multivariate polynomials with multiplicities. This algorithm relies on the reduction to a problem of simultaneous polynomial approximations, which is then solved using fast structured linear algebra techniques. This algorithm leads to the best known complexity bounds for the interpolation step of the list-decoding of Reed-Solomon codes, Parvaresh-Vardy codes or folded Reed-Solomon codes. In the special case of Reed-Solomon codes, it allows to accelerate the interpolation step of Guruswami and Sudan's list-decoding by a factor (list size)/(multiplicity).

6.3.5. Polynomial system solving

M. Bardet (U. Rouen), J.-C. Faugère (PolSys team) and B. Salvy studied the complexity of Gröbner bases computation, in particular in the generic situation where the variables are in simultaneous Noether position with respect to the system. They gave a bound on the number of polynomials of each degree in a Gröbner basis computed by Faugère's F_5 algorithm in this generic case for the grevlex ordering (which is also a bound on the number of polynomials for a reduced Gröbner basis, independently of the algorithm used) and used it to bound the complexity of the F_5 algorithm [5].

6.3.6. Linear differential equations

In [6], A. Bostan (SpecFun team), K. Raschel (U. Tours) and B. Salvy proved that the sequence $(e_n^{\mathfrak{S}})_{n\geq 0}$ of excursions in the quarter plane corresponding to a nonsingular step set $\mathfrak{S} \subseteq \{0, \pm 1\}^2$ with infinite group does not satisfy any nontrivial linear recurrence with polynomial coefficients. Accordingly, in those cases, the trivariate generating function of the numbers of walks with given length and prescribed ending point is not D-finite. Moreover, they displayed the asymptotics of $e_n^{\mathfrak{S}}$. This completes the classification of these walks.

Colleagues from the LAAS (Toulouse) and B. Salvy provided a new method for computing the probability of collision between two spherical space objects involved in a short-term encounter. In this specific framework of conjunction, classical assumptions reduce the probability of collision to the integral of a 2-D normal distribution over a disk shifted from the peak of the corresponding Gaussian function. Both integrand and domain of integration directly depend on the nature of the short-term encounter. Thus the inputs are the combined sphere radius, the mean relative position in the encounter plane at reference time as well as the relative position covariance matrix representing the uncertainties. The method they presented is based on an analytical expression for the integral. It has the form of a convergent power series whose coefficients verify a linear recurrence. It is derived using Laplace transform and properties of D-finite functions. The new method has been intensively tested on a series of test-cases and compares favorably to other existing works [29].

6.4. Lattices and cryptography

6.4.1. Worst-Case to Average-Case Reductions for Module Lattices

Most lattice-based cryptographic schemes are built upon the assumed hardness of the Short Integer Solution (SIS) and Learning With Errors (LWE) problems. Their efficiencies can be drastically improved by switching the hardness assumptions to the more compact Ring-SIS and RingLWE problems. However, this change of hardness assumptions comes along with a possible security weakening: SIS and LWE are known to be at least as hard as standard (worst-case) problems on euclidean lattices, whereas Ring-SIS and Ring-LWE are only known to be as hard as their restrictions to special classes of ideal lattices, corresponding to ideals of some polynomial rings. Adeline Langlois and Damien Stehlé defined the Module-SIS and Module-LWE problems, which bridge SIS with Ring-SIS, and LWE with Ring-LWE, respectively. They proved that these average-case problems are at least as hard as standard lattice problems restricted to module lattices (which themselves bridge arbitrary and ideal lattices). As these new problems enlarge the toolbox of the lattice-based cryptographer, they could prove useful for designing new schemes. Importantly, the worst-case to average-case reductions for the module problems are (qualitatively) sharp, in the sense that there exist converse reductions. This property is not known to hold in the context of Ring-SIS/Ring-LWE: Ideal lattice problems could reveal easy without impacting the hardness of Ring-SIS/Ring-LWE [8].

6.4.2. Semantically Secure Lattice Codes for the Gaussian Wiretap Channel

Cong Ling (Imperial College, UK), Laura Luzzi (ENSEA), Jean-Claude Belfiore (Telecom ParisTech) and Damien Stehlé proposed a new scheme of wiretap lattice coding that achieves semantic security and strong secrecy over the Gaussian wiretap channel. The key tool in their security proof is the flatness factor which characterizes the convergence of the conditional output distributions corresponding to different messages and leads to an upper bound on the information leakage. They not only introduced the notion of secrecy-good lattices, but also proposed the flatness factor as a design criterion of such lattices. Both the modulo-lattice

Gaussian channel and the genuine Gaussian channel are considered. In the latter case, they proposed a novel secrecy coding scheme based on the discrete Gaussian distribution over a lattice, which achieves the secrecy capacity to within a half nat under mild conditions. No a priori distribution of the message is assumed, and no dither is used in their proposed schemes [9].

6.4.3. GGHLite: More Efficient Multilinear Maps from Ideal Lattices

The Garg-Gentry-Halevi (GGH) Graded Encoding Scheme, based on ideal lattices, is the first plausible approximation to a cryptographic multilinear map. Unfortunately, the scheme requires very large parameters to provide security for its underlying encoding re-randomization process. Adeline Langlois, Damien Stehlé and Ron Steinfeld (Monash University, Australia) formalized, simplified and improved the efficiency and the security analysis of the re-randomization process in the GGH construction. This results in a new construction that they called GGHLite. In particular, they first lowered the size of a standard deviation parameter of the GGH re-randomization process from exponential to polynomial in the security parameter. This first improvement is obtained via a finer security analysis of the so-called drowning step of re-randomization, in which they applied the Rényi divergence instead of the conventional statistical distance as a measure of distance between distributions. Their second improvement is to reduce the number of randomizers needed to 2, independently of the dimension of the underlying ideal lattices. These two contributions allowed them to decrease the bit size of the public parameters to $O(\lambda \log^2 \lambda)$ in GGHLite, with respect to the security parameter λ (for a constant multilinearity parameter κ) [22].

6.4.4. LLL reducing with the most significant bits

Let B be a basis of a Euclidean lattice, and $\widetilde{B}$ an approximation thereof. Saruchi (IIT Delhi, India), Ivan Morel, Damien Stehlé and Gilles Villard gave a sufficient condition on the closeness between $\widetilde{B}$ and B so that an LLL-reducing transformation U for $\widetilde{B}$ remains valid for B. Further, they analysed an efficient reduction algorithm when B is itself a small deformation of an LLL-reduced basis. Applications include speeding-up reduction by keeping only the most significant bits of B, reducing a basis that is only approximately known, and efficiently batching LLL reductions for closely related inputs [30].

6.4.5. Hardness of k-LWE and Applications in Traitor Tracing

San Ling (NTU, Singapore), Duong Hieu Phan (LAGA), Damien Stehlé and Ron Steinfeld (Monash University, Australia) introduced the k-LWE problem, a Learning With Errors variant of the k-SIS problem. The Boneh-Freeman reduction from SIS to k-SIS suffers from an exponential loss in k. Ling *et al.* improved and extended it to an LWE to k-LWE reduction with a polynomial loss in k, by relying on a new technique involving trapdoors for random integer kernel lattices. Based on this hardness result, they presented the first algebraic construction of a traitor tracing scheme whose security relies on the worstcase hardness of standard lattice problems. The proposed LWE traitor tracing is almost as efficient as the LWE encryption. Further, it achieves public traceability, i.e., allows the authority to delegate the tracing capability to untrusted parties. To this aim, Ling *et al.* introduced the notion of projective sampling family in which each sampling function is keyed and, with a projection of the key on a well chosen space, one can simulate the sampling function in a computationally indistinguishable way. The construction of a projective sampling family from k-LWE allows us to achieve public traceability, by publishing the projected keys of the users [27].

6.4.6. Lattice-Based Group Signatures Scheme with Verifier-local Revocation

Support of membership revocation is a desirable functionality for any group signature scheme. Among the known revocation approaches, verifier-local revocation (VLR) seems to be the most flexible one, because it only requires the verifiers to possess some up-to-date revocation information, but not the signers. All of the contemporary VLR group signatures operate in the bilinear map setting, and all of them will be insecure once quantum computers become a reality. Adeline Langlois, San Ling, Khoa Nguyen and Huaxiong Wang (NTU, Singapore) introduced the first lattice-based VLR group signature [21], and thus, the first such scheme that is believed to be quantum-resistant. In comparison with existing lattice-based group signatures, this scheme has several noticeable advantages: support of membership revocation, logarithmic-size signatures, and weaker security assumption. In the random oracle model, our scheme is proved to be secure based on the hardness of the Shortest Independent Vector Problem with approximation factor $\gamma = \widetilde{O}(n^{1.5})$ - an assumption that is as weak as those of state-of-the-art lattice-based standard signatures. Moreover, this construction works without relying on encryption schemes, which is an intriguing feature for group signatures.

6.4.7. Proxy Re-Encryption Scheme Supporting a Selection of Delegatees

Julien Devigne (Orange Labs), Eleonora Guerrini (Univ. Montpellier 2, LIRMM) and Fabien Laguillaumie adapt the primitive of proxy re-encryption which allows a user to decide that in case of unavailability, one (or several) particular user, the delegatee, will be able to read his confidential messages. They modify it so that a sender can choose who among many potential delegatees will be able to decrypt his messages, and propose a simple and efficient scheme which is secure under chosen plaintext attack under standard algorithmic assumption in a bilinear setting. They also investigate the possibility to add a traceability of the proxy so that one can detect if it has leaked some re-encryption keys [17].

6.4.8. Practical validation of several fault attacks against the Miller algorithm

Ronan Lashermes (SAS-ENSMSE, PRISM), Marie Paindavoine, Nadia El Mrabet (Univ. P8, LIASD), Jacques Fournier (SAS-ENSMSE) and Louis Goubin (UVSQ, PRISM) describe practical implementations of fault attacks against the Miller algorithm, which computes pairing evaluations on algebraic curves. These implementations validate common fault models used against pairings. In the light of the implemented fault attacks, they show that some blinding techniques proposed to protect the algorithm against Side-Channels Analyses cannot be used as countermeasures against the implemented fault attacks [23].

6.4.9. Non-Malleability from Malleability: Simulation-Sound Quasi-Adaptive NIZK Proofs and CCA2-Secure Encryption from Homomorphic Signatures

Verifiability is central to building protocols and systems with integrity. Initially, efficient methods employed the Fiat-Shamir heuristics. Since 2008, the Groth-Sahai techniques have been the most efficient in constructing non-interactive witness indistinguishable and zero-knowledge proofs for algebraic relations in the standard model. For the important task of proving membership in linear subspaces, Jutla and Roy (Asiacrypt 2013) gave significantly more efficient proofs in the quasi-adaptive setting (QA-NIZK). For membership of the row space of a $t \times n$ matrix, their QA-NIZK proofs save $\Omega(t)$ group elements compared to Groth-Sahai. In [26], Benoît Libert, Thomas Peters (UCL, Belgique), Marc Joye (Technicolor, USA) and Moti Yung (Google and Columbia U, USA) gave QA-NIZK proofs made of a *constant* number group elements – regardless of the number of equations or the number of variables – and additionally proved them *unbounded* simulation-sound. Unlike previous unbounded simulation-sound Groth-Sahai-based proofs, their construction does not involve quadratic pairing product equations and does not rely on a chosen-ciphertext-secure encryption scheme. Instead, they built on structure-preserving signatures with homomorphic properties. They applied their methods to design new and improved CCA2-secure encryption schemes. In particular, they built the first efficient threshold CCA-secure keyed-homomorphic encryption scheme (*i.e.*, where homomorphic operations can only be carried out using a dedicated evaluation key) with publicly verifiable ciphertexts.

6.4.10. Born and Raised Distributively: Fully Distributed Non-Interactive Adaptively-Secure Threshold Signatures with Short Shares

Threshold cryptography is a fundamental distributed computational paradigm for enhancing the availability and the security of cryptographic public-key schemes. It does it by dividing private keys into n shares handed out to distinct servers. In threshold signature schemes, a set of at least $t+1 \leq n$ servers is needed to produce a valid digital signature. Availability is assured by the fact that any subset of $t+1$ servers can produce a signature when authorized. At the same time, the scheme should remain robust (in the fault tolerance sense) and unforgeable (cryptographically) against up to t corrupted servers; *i.e.*, it adds quorum control to traditional cryptographic services and introduces redundancy. Originally, most practical threshold signatures have a number of demerits: They have been analyzed in a static corruption model (where the set of corrupted servers is fixed at the very beginning of the attack), they require interaction, they assume a trusted dealer in the key generation phase (so that the system is not fully distributed), or they suffer from certain overheads in terms of storage (large share sizes).

In [24], Benoît Libert, Marc Joye (Technicolor, USA) and Moti Yung (Google and Columbia U, USA) constructed practical *fully distributed* (the private key is born distributed), non-interactive schemes – where the servers can compute their partial signatures without communication with other servers – with adaptive security (*i.e.*, the adversary corrupts servers dynamically based on its full view of the history of the system). Their schemes are very efficient in terms of computation, communication, and scalable storage (with private key shares of size $O(1)$, where certain solutions incur $O(n)$ storage costs at each server). Unlike other adaptively secure schemes, their schemes are erasure-free (reliable erasure is a hard to assure and hard to administer property in actual systems). Such a fully distributed highly constrained scheme has been an open problem in the area. In particular, and of special interest, is the fact that Pedersen's traditional distributed key generation (DKG) protocol can be safely employed in the initial key generation phase when the system is born – although it is well-known not to ensure uniformly distributed public keys. An advantage of this is that this protocol only takes one round optimistically (in the absence of faulty player).

6.4.11. Concise Multi-challenge CCA-Secure Encryption and Signatures with Almost Tight Security

To gain strong confidence in the security of a public-key scheme, it is most desirable for the security proof to feature a tight reduction between the adversary and the algorithm solving the under-lying hard problem. Recently, Chen and Wee (Crypto '13) described the first Identity-Based Encryption scheme with almost tight security under a standard assumption. Here, "almost tight" means that the security reduction only loses a factor $O(\lambda)$ – where λ is the security parameter – instead of a factor proportional to the number of adversarial queries. Chen and Wee also gave the shortest signatures whose security almost tightly relates to a simple assumption in the standard model. Also recently, Hofheinz and Jager (Crypto '12) constructed the first CCA-secure public-key encryption scheme in the multi-user setting with tight security. These constructions give schemes that are significantly less efficient in length (and thus, processing) when compared with the earlier schemes with loose reductions in their proof of security. Hofheinz and Jager's scheme has a ciphertext of a few hundreds of group elements, and they left open the problem of finding truly efficient constructions. Likewise, Chen and Wee's signatures and IBE schemes are somewhat less efficient than previous constructions with loose reductions from the same assumptions.

In [25], Benoît Libert, Thomas Peters (UCL, Belgique), Marc Joye (Technicolor, USA) and Moti Yung (Google and Columbia U, USA) considered space-efficient schemes with security almost tightly related to standard assumptions. As a step in solving the open question by Hofheinz and Jager, they constructed an efficient CCA-secure public-key encryption scheme whose chosen-ciphertext security in the multi-challenge, multi-user setting almost tightly relates to the DLIN assumption (in the standard model). Quite remarkably, the ciphertext size decreases to 69 group elements under the DLIN assumption whereas the best previous solution required about 400 group elements. Their scheme is obtained by taking advantage of a new almost tightly secure signature scheme (in the standard model) they developed and which is based on the recent concise proofs of linear subspace membership in the quasi-adaptive non-interactive zero-knowledge setting (QA-NIZK) defined by Jutla and Roy (Asiacrypt '13). The new signature scheme reduces the length of the previous such signatures (by Chen and Wee) by 37% under the Decision Linear assumption, by almost 50% under the K-LIN assumption, and it becomes only 3 group elements long under the Symmetric eXternal Diffie-Hellman assumption. Our signatures are obtained by carefully combining the proof technique of Chen and Wee and the above mentioned QA-NIZK proofs.

ASAP Project-Team

6. New Results

6.1. Highlights of the Year

- Anne-Marie Kermarrec is the recipient of the **ACM/IFIP/USENIX/Middleware 10-Years Best Paper Award**, for her paper *The peer sampling service: Experimental evaluation of unstructured gossip-based implementations* (Middleware 2004), co-authored with Márk Jelasity, Rachid Guerraoui, and Maarten van Steen.
- Anne-Marie Kermarrec is the recipient of the **WISE 2014 Best Paper Award**, for her paper [18], co-authored with Alexandra Olteanu and Karl Aberer.
- Michel Raynal is the recipient of the **PODC 2014 Best Paper Award**, for his paper [34], co-authored with Achour Mostefaoui and Moumen Hamouna.
- The MEDIEGO recommendation engine was demonstrated at **Le Web 14** in partnership with FranceTV.

BEST PAPERS AWARDS :

[18] **15th International Conference on Web Information System Engineering (WISE 2014)**. O. ALEXANDRA, A.-M. KERMARREC, K. ABERER.

[34] **ACM PODC**. A. MOSTEFAOUI, M. HAMOUNA, M. RAYNAL.

6.2. Models and abstractions for distributed systems

6.2.1. Signature-free asynchronous Byzantine consensus

Participant: Michel Raynal.

In [34] we present a new round-based asynchronous consensus algorithm that copes with up to $t < n/3$ Byzantine processes, where n is the total number of processes. In addition of not using signature, not assuming a computationally-limited adversary, while being optimal with respect to the value of t, this algorithm has several noteworthy properties: the expected number of rounds to decide is four, each round is composed of two or three communication steps and involves $O(n^2)$ messages, and a message is composed of a round number plus a single bit. To attain this goal, the consensus algorithm relies on a common coin as defined by Rabin, and a new extremely simple and powerful broadcast abstraction suited to binary values. The main target when designing this algorithm was to obtain a cheap and simple algorithm. This was motivated by the fact that, among the first-class properties, simplicity –albeit sometimes under-estimated or even ignored– is a major one.

This is a joint work with Achour Mostéfaouin and Hamouma Moumen. It received the PODC 2014 Best Paper Award.

6.2.2. Randomized mutual exclusion with constant amortized RMR complexity on the DSM

Participant: George Giakkoupis.

In [30] we settle an open question by determining the remote memory reference (RMR) complexity of randomized mutual exclusion, on the distributed shared memory model (DSM) with atomic registers, in a weak but natural (and stronger than oblivious) adversary model. In particular, we present a mutual exclusion algorithm that has constant expected amortized RMR complexity and is deterministically deadlock free. Prior to this work, no randomized algorithm with $o(\log n/\log\log n)$ RMR complexity was known for the DSM model. Our algorithm is fairly simple, and compares favorably with one by Bender and Gilbert (FOCS 2011) for the CC model, which has expected amortized RMR complexity $O(\log^2 \log n)$ and provides only probabilistic deadlock freedom.

This is a joint work with Philipp Woelfel (Univ. of Calgary, Canada).

6.2.3. Reliable shared memory abstraction on top of asynchronous Byzantine message-passing systems

Participants: Michel Raynal, Julien Stainer.

This work is on the construction and the use of a shared memory abstraction on top of an asynchronous message-passing system in which up to t processes may commit Byzantine failures. This abstraction consists of arrays of n single-writer/multi-reader atomic registers, where n is the number of processes. Differently from usual atomic registers which record a single value, each of these atomic registers records the whole history of values written to it. A distributed algorithm building such a shared memory abstraction it first presented. This algorithm assumes $t < n/3$, which is shown to be a necessary and sufficient condition for such a construction. Hence, the algorithm is resilient-optimal. Then we present distributed algorithms built on top of this shared memory abstraction, which cope with up to t Byzantine processes. The simplicity of these algorithms constitutes a strong motivation for such a shared memory abstraction in the presence of Byzantine processes. For a lot of problems, algorithms are more difficult to design and prove correct in a message-passing system than in a shared memory system. Using a protocol stacking methodology, the aim of the proposed abstraction is to allow an easier design (and proof) of distributed algorithms, when the underlying system is an asynchronous message-passing system prone to Byzantine failures.

This work was done in collaboration with Damien Imbs and Sergio Rajsbaum. It has been published in SIRROCCO [32] and as a technical report [43].

6.2.4. Distributed Universality

Participants: Michel Raynal, Julien Stainer.

A notion of a universal construction suited to distributed computing has been introduced by M. Herlihy in his celebrated paper "Wait-free synchronization" (ACM TOPLAS, 1991). A universal construction is an algorithm that can be used to wait-free implement any object defined by a sequential specification. Herlihy's paper shows that the basic system model, which supports only atomic read/write registers, has to be enriched with consensus objects to allow the design of universal constructions. The generalized notion of a k-universal construction has been recently introduced by Gafni and Guerraoui (CONCUR 2011). A k-universal construction is an algorithm that can be used to simultaneously implement k objects (instead of just one object), with the guarantee that at least one of the k constructed objects progresses forever. While Herlihy's universal construction relies on atomic registers and consensus objects, a k-universal construction relies on atomic registers and k-simultaneous consensus objects (which are wait-free equivalent to k-set agreement objects in the read/write system model). This work significantly extends the universality results introduced by Herlihy and Gafni-Guerraoui. In particular, we present a k-universal construction which satisfies the following five desired properties, which are not satisfied by the previous k-universal construction: (1) among the k objects that are constructed, at least ℓ objects (and not just one) are guaranteed to progress forever; (2) the progress condition for processes is wait-freedom, which means that each correct process executes an infinite number of operations on each object that progresses forever; (3) if any of the k constructed objects stops progressing, all its copies (one at each process) stop in the same state; (4) the proposed construction is contention-aware, in the sense that it uses only read/write registers in the absence of contention; and (5) it is generous with respect to the obstruction-freedom progress condition, which means that each process is able to complete any one of its pending operations on the k objects if all the other processes hold still long enough. The proposed construction, which is based on new design principles, is called a (k,ℓ)-universal construction. It uses a natural extension of k-simultaneous consensus objects, called (k,ℓ)-simultaneous consensus objects ((k,ℓ)-SC). Together with atomic registers, (k,ℓ)-SC objects are shown to be necessary and sufficient for building a (k,ℓ)-universal construction, and, in that sense, (k,ℓ)-SC objects are (k,ℓ)-universal.

This work was done in collaboration with Gadi Taubenfeld. It has been published as a brief announcement in PODC [37] and the full version appeared in OPODIS [38]. A version has also been published as a technical report [45].

6.2.5. Computing in the presence of concurrent solo executions

Participants: Michel Raynal, Julien Stainer.

In a wait-free model any number of processes may crash. A process runs solo when it computes its local output without receiving any information from other processes, either because they crashed or they are too slow. While in wait-free shared-memory models at most one process may run solo in an execution, any number of processes may have to run solo in an asynchronous wait-free message-passing model. This work is on the computability power of models in which several processes may concurrently run solo. It first introduces a family of round-based wait-free models, called the d-solo models, $1 \leq d \leq n$, where up to d processes may run solo. We then give a characterization of the colorless tasks that can be solved in each d-solo model. We also introduce the (d,ϵ)-solo approximate agreement task, which generalizes ϵ-approximate agreement, and proves that (d,ϵ)-solo approximate agreement can be solved in the d-solo model, but cannot be solved in the $(d+1)$-solo model. We study also the relation linking d-set agreement and (d,ϵ)-solo approximate agreement in asynchronous wait-free message-passing systems. These results establish for the first time a hierarchy of wait-free models that, while weaker than the basic read/write model, are nevertheless strong enough to solve non-trivial tasks.

This work was done in collaboration with Maurice Herlihy and Sergio Rajsbaum. It has been published in LATIN [31].

6.2.6. A simple broadcast algorithm for recurrent dynamic systems

Participants: Michel Raynal, Julien Stainer.

This work presents a simple broadcast algorithm suited to dynamic systems where links can repeatedly appear and disappear. The algorithm is proved correct and a simple improvement is introduced, that reduces the number and the size of control messages. As it extends in a simple way a classical network traversal algorithm to the dynamic context, the proposed algorithm has also pedagogical flavor.

This work was done in collaboration with Jiannong Cao and Weigang Wu. It has been published in AINA [36].

6.2.7. Fisheye consistency: Keeping data in synch in a georeplicated world

Participants: Michel Raynal, François Taïani.

Over the last thirty years, numerous consistency conditions for replicated data have been proposed and implemented. Popular examples of such conditions include linearizability (or atomicity), sequential consistency, causal consistency, and eventual consistency. These consistency conditions are usually defined independently from the computing entities (nodes) that manipulate the replicated data; i.e., they do not take into account how computing entities might be linked to one another, or geographically distributed. To address this lack, as a first contribution, this work [41] introduces the notion of proximity graph between computing nodes. If two nodes are connected in this graph, their operations must satisfy a strong consistency condition, while the operations invoked by other nodes are allowed to satisfy a weaker condition. The second contribution is the use of such a graph to provide a generic approach to the hybridization of data consistency conditions into the same system. We illustrate this approach on sequential consistency and causal consistency, and present a model in which all data operations are causally consistent, while operations by neighboring processes in the proximity graph are sequentially consistent. The third contribution of this work is the design and the proof of a distributed algorithm based on this proximity graph, which combines sequential consistency and causal consistency (the resulting condition is called fisheye consistency). In doing so this work not only extends the domain of consistency conditions, but provides a generic provably correct solution of direct relevance to modern georeplicated systems.

This work was done in collaboration with Roy Friedman (The Technion, Haifa, Israel)

6.3. Large-scale and user-centric distributed systems

6.3.1. Archiving cold data in warehouses with clustered network coding

Participants: Fabien André, Anne-Marie Kermarrec.

Modern storage systems now typically combine plain replication and erasure codes to reliably store large amount of data in datacenters. Plain replication allows a fast access to popular data, while erasure codes, e.g., Reed-Solomon codes, provide a storage-efficient alternative for archiving less popular data. Although erasure codes are now increasingly employed in real systems, they experience high overhead during maintenance, i.e., upon failures, typically requiring files to be decoded before being encoded again to repair the encoded blocks stored at the faulty node.

In this work, we proposed a novel erasure code system, tailored for networked archival systems. The efficiency of our approach relies on the joint use of random codes and a clustered placement strategy. Our repair protocol leverages network coding techniques to reduce by 50% the amount of data transferred during maintenance, by repairing several cluster files simultaneously. We demonstrated both through an analysis and extensive experimental study conducted on a public testbed that our approach significantly decreases both the bandwidth overhead during the maintenance process and the time to repair lost data. We also showed that using a non-systematic code does not impact the throughput, and comes only at the price of a higher CPU usage. Based on these results, we evaluated the impact of this higher CPU consumption on different configurations of data coldness by determining whether the cluster's network bandwidth dedicated to repair or CPU dedicated to decoding saturates first.

This work has been conducted in collaboration with Erwan Le Merrer, Nicolas Le Scouarnec, Gilles Straub (Technicolor) and A. van Kempen (Univ. Nantes) and published in ACM Eurosys 2014 [19].

6.3.2. WebGC: Browser-based gossiping

Participants: Raziel Carvajal Gomez, Davide Frey, Anne-Marie Kermarrec.

The advent of browser-to-browser communication technologies like WebRTC has renewed interest in the peer-to-peer communication model. However, the available WebRTC code base still lacks important components at the basis of several peer-to-peer solutions. Through a collaboration with Mathieu Simonin from the Inria SED in the context of the Brow2Brow ADT project, we started to tackle this problem by proposing WebGC, a library for gossip-based communication between web browsers. Due to their inherent scalability, gossip-based, or epidemic protocols constitute a key component of a large number of decentralized applications. WebGC thus represents an important step towards their wider spread. We demonstrated a preliminary version of the library at Middleware 2014 [47].

6.3.3. Large-scale graph processing in datacenters with bandwidth guarantees

Participants: Nitin Chiluka, Anne-Marie Kermarrec.

Recent research has shown that the performance of data-intensive applications in multi-tenant datacenters can be severely impacted by each other's network usage. Starvation for network bandwidth in such datacenters typically results in significantly longer completion times for large-scale distributed applications. To address this concern, researchers propose bandwidth guarantees for all the virtual machines (VMs) initiated by each tenant in the datacenter in order to provide a predictable performance for their applications. In our work, we focus on large-scale graph processing in such datacenters. More specifically, given k VMs with their respective bandwidth constraints and a large graph, we perform a k-way partition on the graph such that the subsequent computation of various algorithms (e.g., PageRank, graph factorization) take minimal time.

6.3.4. Scaling KNN computation over large graphs on a PC

Participants: Nitin Chiluka, Anne-Marie Kermarrec, Javier Olivares.

Frameworks such as GraphChi and X-Stream are increasingly gaining attention for their ability to perform scalable computation on large graphs by leveraging disk and memory on a single commodity PC. These frameworks rely on the graph structure to remain the same for the entire pe- riod of computation of various algorithms such as PageRank and triangle counting. As a consequence, these frameworks are not applicable to algorithms that require the graph struc- ture to change during their computation. In this work, we focus on one such algorithm – K-Nearest Neighbors (KNN) – which is widely used in recommender systems. Our approach aims to minimize random accesses to disk as well as the amount of data loaded/unloaded from/to disk so as to better utilize the computational power, thus improving the algorithmic efficiency. The preliminary design and results of our approach appeared in Middleware 2014 [23].

6.3.5. Privacy-preserving distributed collaborative filtering

Participants: Davide Frey, Arnaud Jégou, Anne-Marie Kermarrec.

In collaboration with Antoine Boutet from the Univ. St Etienne, and Rachid Guerraoui from EPFL, we proposed a new mechanism to preserve privacy while leveraging user profiles in distributed recommender systems. Our mechanism relies on two contributions: (*i*) an original obfuscation scheme, and (*ii*) a randomized dissemination protocol. We showed that our obfuscation scheme hides the exact profiles of users without significantly decreasing their utility for recommendation. In addition, we precisely characterized the conditions that make our randomized dissemination protocol differentially private.

We compared our mechanism with a non-private as well as with a fully private alternative. We considered a real dataset from a user survey and report on simulations as well as planetlab experiments. In short, our extensive evaluation showed that our twofold mechanism provides a good trade-off between privacy and accuracy, with little overhead and high resilience.

6.3.6. Behave: Behavioral cache for web content

Participants: Davide Frey, Anne-Marie Kermarrec.

In collaboration with Mathieu Goessens, a former intern of the team, we proposed Behave: a novel approach for peer-to-peer cache-oriented applications such as CDNs. Behave relies on the principle of Behavioral Locality inspired from collaborative filtering. Users that have visited similar websites in the past will have local caches that provide interesting content for one another.

Behave exploits epidemic protocols to build overlapping communities of peers with similar interests. Peers in the same one-hop community federate their cache indexes in a Behavioral cache. Extensive simulations on a real data trace show that Behave can provide zero-hop lookup latency for about 50% of the content available in a DHT-based CDN. The results of this work were published at DAIS 2014 [26].

6.3.7. HyRec: Leveraging browsers for scalable recommenders

Participants: Davide Frey, Anne-Marie Kermarrec.

The ever-growing amount of data available on the Internet calls for personalization. Yet, the most effective personalization schemes, such as those based on collaborative filtering (CF), are notoriously resource greedy. In this work, we proposed HyRec, an online cost-effective scalable system for user-based CF personalization. HyRec offloads recommendation tasks onto the web browsers of users, while a server orchestrates the process and manages the relationships between user profiles.

We fully implemented HyRec and we extensively evaluated it on several workloads from MovieLens and Digg. Our experiments conveyed the ability of HyRec to reduce the operation costs of content providers by nearly 50% and to provide a 100-fold improvement in scalability with respect to a centralized (or cloud-based recommender approach), while preserving the quality of personalization. HyRec is also virtually transparent to users and induces only 3% of the bandwidth consumption of a p2p solution. This work was done in collaboration with Antoine Boutet from the Univ. St Etienne, as well as with Rachid Guerraoui, and Rhicheek Patra from EPFL. It resulted in a publication at Middleware 2014 [22].

6.3.8. Landmark-based similarity for p2p collaborative filtering

Participants: Davide Frey, Anne-Marie Kermarrec, Antoine Rault, François Taïani.

Computing k-nearest-neighbor graphs constitutes a fundamental operation in a variety of data-mining applications. As a prominent example, user-based collaborative-filtering provides recommendations by identifying the items appreciated by the closest neighbors of a target user. As this kind of applications evolve, they will require KNN algorithms to operate on more and more sensitive data. This has prompted researchers to propose decentralized peer-to-peer KNN solutions that avoid concentrating all information in the hands of one central organization. Unfortunately, such decentralized solutions remain vulnerable to malicious peers that attempt to collect and exploit information on participating users.

We seek to overcome this limitation by proposing *H&S* (Hide & Share), a novel landmark-based similarity mechanism for decentralized KNN computation. Landmarks allow users (and the associated peers) to estimate how close they lay to one another without disclosing their individual profiles.

We evaluate *H&S* in the context of a user-based collaborative-filtering recommender with publicly available traces from existing recommendation systems. We show that although landmark-based similarity does disturb similarity values (to ensure privacy), the quality of the recommendations is not as significantly hampered. We also show that the mere fact of disturbing similarity values turns out to be an asset because it prevents a malicious user from performing a profile reconstruction attack against other users, thus reinforcing users' privacy. Finally, we provide a formal privacy guarantee by computing the expected amount of information revealed by *H&S* about a user's profile.

This work was done in collaboration with Jingjing Wang, and Rachid Guerraoui.

6.3.9. Adaptation for the masses: Towards decentralized adaptation in large-scale p2p recommenders

Participants: Davide Frey, Anne-Marie Kermarrec, François Taïani.

Decentralized recommenders have been proposed to deliver privacy-preserving, personalized and highly scalable on-line recommendation services. Current implementations tend, however, to rely on hard-wired, mechanisms that cannot adapt. Deciding beforehand which hard-wired mechanism to use can be difficult, as the optimal choice might depend on conditions that are unknown at design time. In [27], we have proposed a framework to develop dynamically adaptive decentralized recommendation systems. Our proposal supports a decentralized form of adaptation, in which individual nodes can independently select, and update their own recommendation algorithm, while still collectively contributing to the overall system's services.

This work was done in collaboration with Christopher Maddock and Andreas Mauthe (Univ. of Lancaster, UK).

6.3.10. Tight bounds for rumor spreading with vertex expansion

Participant: George Giakkoupis.

In [28] we establish an upper bound for the classic PUSH-PULL rumor spreading protocol on general graphs, in terms of the vertex expansion of the graph. We show that $O(\log^2(n)/\alpha)$ rounds suffice with high probability to spread a rumor from any single node to all n nodes, in any graph with vertex expansion at least α. This bound matches a known lower bound, and settles the natural question on the relationship between rumor spreading and vertex expansion asked by Chierichetti, Lattanzi, and Panconesi (SODA 2010). Further, some of the arguments used in the proof may be of independent interest, as they give new insights, for example, on how to choose a small set of nodes in which to plant the rumor initially, to guarantee fast rumor spreading.

6.3.11. Greedy routing in small-world networks with power-law degrees

Participant: George Giakkoupis.

In [12] we study decentralized routing in small-world networks that combine a wide variation in node degrees with a notion of spatial embedding. Specifically, we consider a variant of J. Kleinberg's grid-based small-world model in which (1) the number of long-range edges of each node is not fixed, but is drawn from a power-law probability distribution with exponent parameter $\alpha \geq 0$ and constant mean, and (2) the long-range edges are considered to be bidirectional for the purposes of routing. This model is motivated by empirical observations indicating that several real networks have degrees that follow a power-law distribution. The measured power-law exponent α for these networks is often in the range between 2 and 3. For the small-world model we consider, we show that when $2 < \alpha < 3$ the standard greedy routing algorithm, in which a node forwards the message to its neighbor that is closest to the target in the grid, finishes in an expected number of $O(\log^{\alpha-1} n \cdot \log\log n)$ steps, for any source–target pair. This is asymptotically smaller than the $O(\log^2 n)$ steps needed in Kleinberg's original model with the same average degree, and approaches $O(\log n)$ as α approaches 2. Further, we show that when $0 \leq \alpha < 2$ or $\alpha \geq 3$ the expected number of steps is $O(\log^2 n)$, while for $\alpha = 2$ it is $O(\log^{4/3} n)$. We complement these results with lower bounds that match the upper bounds within at most a $\log\log n$ factor.

This is a joint work with Pierre Fraigniaud (Inria Paris-Rocquencourt and CNRS).

6.3.12. Randomized rumor spreading in dynamic graphs

Participant: George Giakkoupis.

In [29] we consider the well-studied rumor spreading model in which nodes contact a random neighbor in each round in order to push or pull the rumor. Unlike most previous works which focus on static topologies, we look at a dynamic graph model where an adversary is allowed to rewire the connections between vertices before each round, giving rise to a sequence of graphs, $G_1, G_2, ...$ Our first result is a bound on the rumor spreading time in terms of the conductance of those graphs. We show that if the degree of each node does not change much during the protocol (that is, by at most a constant factor), then the spread completes within t rounds for some t such that the sum of conductances of the graphs G_1 up to G_t is $O(\log n)$. This result holds even against an adaptive adversary whose decisions in a round may depend on the set of informed vertices before the round, and implies the known tight bound with conductance for static graphs. Next we show that for the alternative expansion measure of vertex expansion, the situation is different. An adaptive adversary can delay the spread of rumor significantly even if graphs are regular and have high expansion, unlike in the static graph case where high expansion is known to guarantee fast rumor spreading. However, if the adversary is oblivious, i.e., the graph sequence is decided before the protocol begins, then we show that a bound close to the one for the static case holds for any sequence of regular graphs.

This is a joint work with Thomas Sauerwald (Univ. of Cambridge, UK) and Alexandre Stauffer (Univ. of Bath, UK).

6.3.13. Privacy-preserving dissemination in social networks and microblogs

Participants: George Giakkoupis, Arnaud Jégou, Anne-Marie Kermarrec, Nupur Mittal.

Online micro-blogging services and social networks, as exemplified by Twitter and Facebook, have emerged as an important means of disseminating information quickly and at large scale. A standard mechanism in micro-blogging that allows for interesting content to reach a wider audience is that of *reposting* (i.e., *retweeting* in Twitter, or *sharing* in Facebook) of content initially posted by another user. Motivated by recent events in which users were prosecuted merely for reposting anti-government information, we present in [42] Riposte, a randomized reposting scheme that provides privacy guarantees against such charges. The idea is that if the user likes a post, Riposte will repost it only with some (carefully chosen) probability; and if the user does not like it, Riposte may still repost it with a slightly smaller probability. These probabilities are computed for each user as a function of the number of connections of the user in the network, and the extent to which the post has already reached those connections. The choice of these probabilities is based on results for branching processes, and ensures that interesting posts (liked by a large fraction of users) are likely to disseminate widely, whereas uninteresting posts (or spam) do not spread. Riposte is executed locally at the user, thus the user's opinion on the post is not communicated to the micro-blogging server. In this work, we quantify Riposte's ability to protect users in terms of differential privacy and provide analytical bounds on the dissemination of posts. We also do extensive experiments based on topologies of real networks, including Twitter, Facebook, Renren, Google+ and LiveJournal.

This work has been carried out in collaboration with Rachid Guerraoui (EPFL).

6.3.14. Adaptive streaming

Participants: Ali Gouta, Anne-Marie Kermarrec.

HTTP Adaptive Streaming (HAS) is gradually being adopted by Over The Top (OTT) content providers. In HAS, a wide range of video bitrates of the same video content are made available over the internet so that clients' players pick the video bitrate that best fit their bandwidth. Yet, this affects the performance of some major components of the video delivery chain, namely CDNs or transparent caches since several versions of the same content compete to be cached. In this context, we investigated the benefits of a Cache Friendly HAS system (CF-DASH), which aims to improve the caching efficiency in mobile networks and to sustain the quality of experience of mobile clients. We conducted our work by presenting a set of observations we made on a large number of clients requesting HAS contents. We introduced the CF-Dash system and our testbed implementation. Finally, we evaluated CF-dash based on trace-driven simulations and testbed experiments. Our validation results are promising. Simulations on real HAS traffic show that we achieve a significant gain in hit-ratio that ranges from 15% up to 50%. This work was done in collaboration with Zied Aouini, Yannick Le Louedec and Diallo Mamadou, and was published in NOSSDAV 2014 [39].

6.3.15. Predictive capabilities of social and interest affinity for recommendations

Participant: Anne-Marie Kermarrec.

The advent of online social networks created new prediction opportunities for recommender systems: instead of relying on past rating history through the use of collaborative filtering (CF), they can leverage the social relations among users as a predictor of user tastes similarity. Alas, little effort has been put into understanding when and why (e.g., for which users and what items) the social affinity (i.e., how well connected users are in the social network) is a better predictor of user preferences than the interest affinity among them as algorithmically determined by CF, and how to better evaluate recommendations depending on, for instance, what type of users a recommendation application targets. This overlook is explained in part by the lack of a systematic collection of datasets including both the explicit social network among users and the collaborative annotated items. In this work, we conducted an extensive empirical analysis on six real-world publicly available datasets, which dissects the impact of user and item attributes, such as the density of social ties or item rating patterns, on the performance of recommendation strategies relying on either the social ties or past rating similarity. Our findings represent practical guidelines that can assist in future deployments and mixing schemes. This work has been done in collaboration with Karl Aberer and Alexandra Olteanu (EPFL Swizerland). The paper received the Best Paper Award at the WISE International Conference [18].

6.3.16. Polystyrene: The decentralized data shape that never dies

Participants: Anne-Marie Kermarrec, François Taïani.

Decentralized topology construction protocols organize nodes along a predefined topology (e.g. a torus, ring, or hypercube). Such topologies have been used in many contexts ranging from routing and storage systems, to publish-subscribe and event dissemination. Since most topologies assume no correlation between the physical location of nodes and their positions in the topology, they do not handle catastrophic failures well, in which a whole region of the topology disappears. When this occurs, the overall shape of the system typically gets lost. This is highly problematic in applications in which overlay nodes are used to map a virtual data space, be it for routing, indexing or storage. In this work [20], we propose a novel decentralized approach that maintains the initial shape of the topology even if a large (consecutive) portion of the topology fails. Our approach relies on the dynamic decoupling between physical nodes and virtual ones enabling a fast reshaping. For instance, our results show that a 51,200-node torus converges back to a full torus in only 10 rounds after 50% of the nodes have crashed. Our protocol is both simple and flexible and provides a novel form of collective survivability that goes beyond the current state of the art.

This work has been done in collaboration with Simon Bouget (ENS Rennes) and Hoel Kervadec (INSA Rennes).

6.3.17. Link-prediction for very large scale graphs using distributed graph engines

Participants: Anne-Marie Kermarrec, François Taïani, Juan Manuel Tirado Martin.

In this project, we consider how the emblematic problem of link-prediction can be implemented efficiently in gather-apply-scatter (GAS) platforms, a popular distributed graph-computation model. Our proposal, called SNAPLE, exploits a novel highly-localized vertex scoring technique, and minimizes the cost of data flow while maintaining prediction quality. When used within GraphLab, SNAPLE can scale to extremely large graphs that a standard implementation of link prediction on cannot handle within the same platform. More precisely, we show that our approach can process a graph containing 1.4 billions edges on a 256 cores cluster in less than three minutes, with no penalty in the quality of predictions. This result corresponds to an over-linear speedup of 30 against a 20-core stand-alone machine running a non-distributed state-of-the-art solution.

6.3.18. GOSSIPKIT: A unified component framework for gossip

Participant: François Taïani.

Although the principles of gossip protocols are relatively easy to grasp, their variety can make their design and evaluation highly time consuming. This problem is compounded by the lack of a unified programming framework for gossip, which means developers cannot easily reuse, compose, or adapt existing solutions to fit their needs, and have limited opportunities to share knowledge and ideas. In [17], we have considered how component frameworks, which have been widely applied to implement middleware solutions, can facilitate the development of gossip-based systems in a way that is both generic and simple. We show how such an approach can maximise code reuse, simplify the implementation of gossip protocols, and facilitate dynamic evolution and re-deployment.

This work was done in collaboration with Shen Lin (SAP Labs) and Gordon Blair (Univ. of Lancaster, UK).

6.3.19. Towards a new model for cyber foraging

Participant: François Taïani.

Cyber foraging seeks to expand the capabilities and battery life of mobile devices by offloading intensive computations to nearby computing nodes (the surrogates). Although promising, current approaches to cyber foraging tend to impose a strict separation between the application state maintained on the mobile device, and data processed on the surrogates. In [33], we argue that this separation limits the applicability of cyber foraging, and explore how state sharing could be implemented in practice.

This work was done in collaboration with Diogo Lima and Hugo Miranda (Univ. of Lisbon, Portugal).

ASCLEPIOS Project-Team

5. New Results

5.1. Highlights of the Year

- Nicholas Ayache was elected a member of the Académie des sciences on 18th Nov. 2014.
- Nicholas Ayache received the “Grand Prix Inria – Académie des sciences 2014“ for his major contributions to Informatics and Computational Sciences at Inria.
- Nicholas Ayache taught the "Personalized Digital Patient" course at the Collège de France on the annual chair "Informatics and Computational Sciences".
- Hervé Lombaert was awarded and ranked 1st in computer science at the highly selective NSERC Postdoctoral Fellowship (Top funding agency in Canada).
- Nina Miolane and Bishesh Khanal won the first prize in the “Popular Vote Awards” at the MICCAI 2014 Educational Challenge for their video on “Statistics on Lie groups for Computational Anatomy“.

BEST PAPER AWARD :

[12] **MICCAI Workshop on Abdominal Imaging – Computational and Clinical Applications**. C. AUDIGIER, T. MANSI, H. DELINGETTE, S. RAPAKA, V. MIHALEF, D. CARNEGIE, E. BOCTOR, M. CHOTI, A. KAMEN, D. COMANICIU, N. AYACHE.

5.2. Medical Image Analysis

5.2.1. 3D/2D Coronary Arteries Registration

Participants: Thomas Benseghir [correspondent], Grégoire Malandain [Morpheme Team], Régis Vaillant [GE-Healthcare], Nicholas Ayache.

This work has been performed in colaboration with GE-Healthcare (Buc) and the Morpheme team at Inria SAM.

3D/2D Registration, Computed Tomography Angiography, X-ray Fluoroscopy, Coronary Arteries, Vascular Tree

Integrating vessel calcifications and occlusion information, extracted from pre-operative 3D CT angiography images into a live fluoroscopic 2D image can greatly improve the guidance of percutaneous coronary interventions. Such task requires a registration step that must provide relevant correspondences between these two complementary modalities. We are developing a framework aiming at preserving the topology of the vascular structures matched between both images.

The introduction of topology in the pairing procedure allows to decrease the mismatches with respect to geometrically-based pairing procedures (e.g. Iterative Closest Point), which, in turn, improves the success rate of the registration method. This is exemplified by Fig. 3 where the proposed pairing method is compared to ICP.

5.2.2. Video Synchronization: An Approach Towards Endoscopic Re-localization

Participants: Anant Suraj Vemuri [correspondent], Nicholas Ayache.

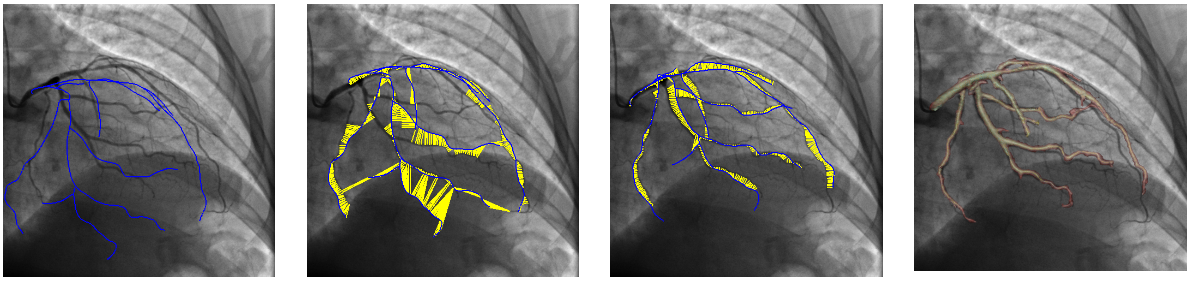


Figure 3. From left to right: 1. initial pose estimate (3D centerlines projection in blue); 2. iterative closest point algorithm registered position (point-pairings in yellow); 3: proposed method registered position; 4: resulting fusion between the two modalities

Endoscopy, Barrett's Esophagus, Re-localization, Electromagnetic Tracking

- Barrett's esophagus is the pre-malignant lesion for the majority of patients with esophageal adenocarcinoma. The evolution of the disease involves endoscopic surveillance for patients every 3-6 months, according to the Seattle protocol.
- The approach is labor-intensive and the primary problem is the inter-operative re-localization of these biopsy sites to guide the treatment.
- In an earlier work we had proposed a general framework for inter-operative biopsy site re-localization framework by introducing an Electro-magnetic tracking system (EMTS) into the loop and providing a way to inter-operatively register video sequences to provide a guided navigation in the esophagus.
- This work has been extended further to fit the operating room workflow. Two external landmarks have been added to the system setup as shown in Fig.4 , to obtain a complete reference frame with respect to the patient and to make the registration, patient specific [38]. The patient-localized reference frame allows the recovery of complete $SE(3)$ including the roll angle about the esophageal axis.

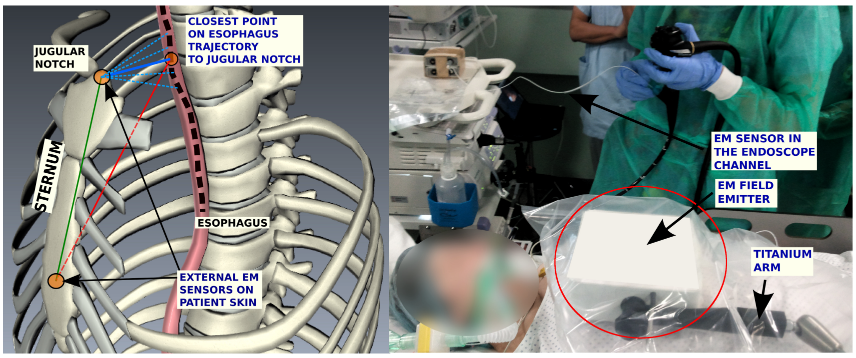


Figure 4. (Left) The orange circular markings indicate the position of the sensors taken as anatomical landmarks to form the reference frame for the patient. (Right) System setup in the operating room.

5.2.3. A sparse Bayesian framework for non-rigid registration

Participants: Loic Le Folgoc [correspondent], Hervé Delingette, Antonio Criminisi, Nicholas Ayache.

This work has been partly supported by Microsoft Research - Inria joint laboratory through its PhD Scholarship Programme and the European Research Council through the ERC Advanced Grant MedYMA (on Biophysical Modeling and Analysis of Dynamic Medical Images).

Registration, Automatic Relevance Determination, Uncertainty Quantification

We propose a sparse Bayesian framework for non-rigid registration. It provides a principled approach to efficiently find an optimal, sparse parameterization of deformations among any preset, widely overcomplete range of basis functions. It addresses open challenges in state-of-the-art registration, such as the automatic joint estimate of model parameters (e.g. noise and regularization levels). We have evaluated the feasibility and performance of our approach on cine MR, tagged MR and 3D US cardiac images, and show state-of-the-art results on benchmark datasets evaluating accuracy of motion and strain (see Fig.5). This work was presented during the MICCAI 2014 conference[20].

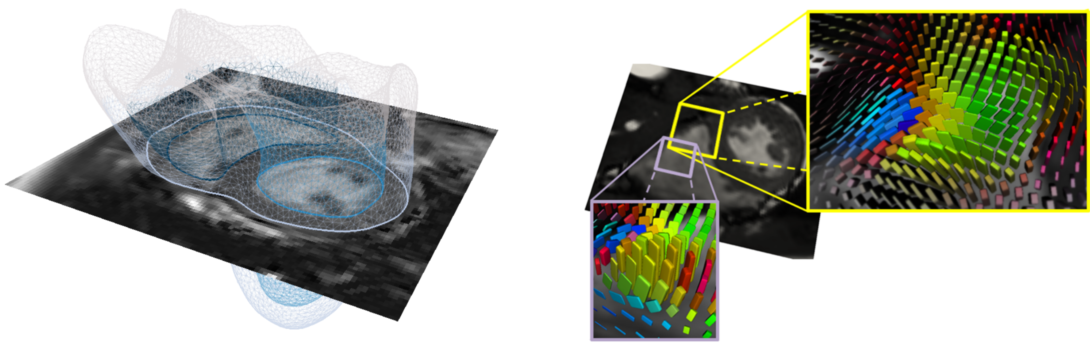


Figure 5. (Left) Mesh contour propagated to end systole via the registration output; (Right) Spatial uncertainty visualized as a tensor map.

5.2.4. Segmentation and anatomic variability of the cochlea and other temporal bone structures from medical images

Participants: Thomas Demarcy [correspondent], Hervé Delingette, Clair Vandersteen, Dan Gnansia [Oticon Medical], Nicholas Ayache.

This work is funded by a CIFRE grant involving Oticon Medical (Vallauris) and performed in collaboration with the IUFC Nice (Pr. Guevara) and CHU Nice (Pr. Raffaelli) .

image segmentation ; surgery planning ; shape modelling ; anatomic variability ; cochlear implant ; temporal bone

- We applied semi-automatic segmentation methods to extract anatomical structures on the inner ear on both micro-CT and CT scan images.
- μ-CT and CT images acquired on the same subject were fused with their segmentation.
- We designed a teaching tool[37] for advanced visualization of temporal bone structures (see Fig. 6).

5.2.5. Understanding cardiac planes of acquisition

Participants: Jan Margeta [correspondent], Nicholas Ayache, Daniel C Lee [Northwestern University], Antonio Criminisi [Microsoft Research Cambridge].

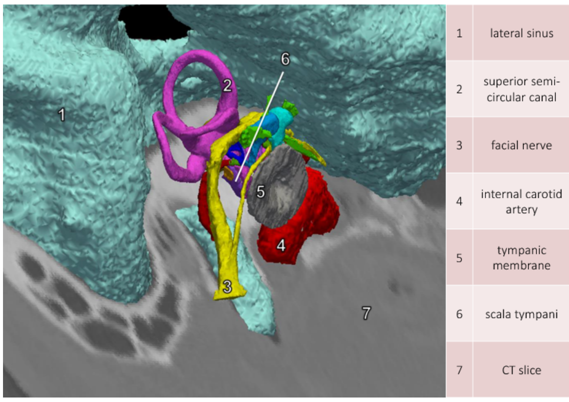


Figure 6. Virtual view of the posterior tympanotomy approach fused with a CT-scan

This work has been partly supported by Microsoft Research through its PhD Scholarship Programme, by ERC Advanced Grant MedYMA (on Biophysical Modeling and Analysis of Dynamic Medical Images), and by the VP2HF FP7 research project.

Cardiac imaging, Machine learning, Magnetic resonance, Data wrangling

DICOM image format defines several tags by which the images can be queried and filtered. Many useful tags are however not standardized and must be cleaned prior to any large scale analysis.

- We developed a machine learning method for automatic recognition of cardiac planes of acquisition (See Fig. 7 for sample predictions).
- Our image based method achieved state of the art performance.
- This work was presented[23] at the Medical Image Understanding and Analysis conference in London.

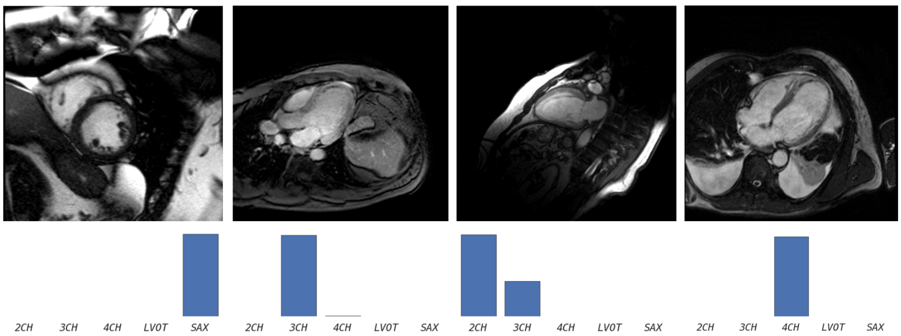


Figure 7. Examples of cardiac acquisition plane predictions and confidences

5.3. Computational Anatomy

5.3.1. Statistical Analysis of Diffusion Tensor Images of the Brain

Participants: Marco Lorenzi [Correspondent], Nicholas Ayache, Xavier Pennec.

Image non-linear registration, Longitudinal modeling, Alzheimer's disease

Alzheimer's disease is characterized by the co-occurrence of different phenomena, starting from the deposition of amyloid plaques and neurofibrillary tangles, to the progressive synaptic, neuronal and axonal damages. The brain atrophy is a sensitive marker of disease progression from pre-clinical to the pathological stages, and computational methods for the analysis of magnetic resonance images of the brain are currently used for group-wise (cross-sectional) and longitudinal studies of pathological morphological changes in clinical populations. The aim of this project is to develop robust and effective computational instruments for the analysis of longitudinal brain changes. In particular novel methods based on non-linear diffeomorphic registration have been investigated in order to reliably detect and statistically analyze pathological morphological changes [5] (see Fig.8). This project is also focused in the comparison of the trajectories of longitudinal morphological changes [31] estimated in different patients. This is a central topic for the development of statistical atlases of the longitudinal evolution of brain atrophy.

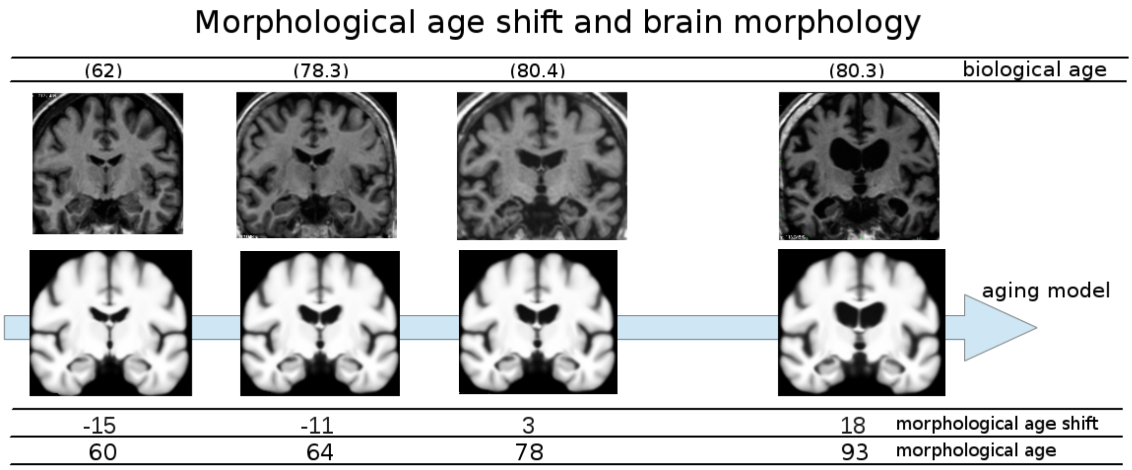


Figure 8. Modeled longitudinal brain changes in normal aging extrapolated from -15 to 18 years, and corresponding observed patient anatomies with estimated morphological age and age shift (biological age in parenthesis). Our modeling framework describes meaningful anatomical changes observed in clinical groups.

5.3.2. Statistical Learning via Synthesis of Medical Images

Participants: Hervé Lombaert [Correspondent], Nicholas Ayache, Antonio Criminisi.

This work has been partly supported by a grant from Microsoft Research-Inria Joint Centre, by ERC Advanced Grant MedYMA (on Biophysical Modeling and Analysis of Dynamic Medical Images)

Statistical learning, Synthesis

Machine learning approaches typically require large training datasets in order to capture as much variability as possible. Application of conventional learning methods on medical images is difficult due to the large variability that exists among patients, pathologies, and image acquisitions. The project aims at exploring how realistic image synthesis could be used, and improve existing machine learning methods.

First year tackled the problem of better exploiting existing training sets, via a smart modeling of the image space (Fig. 9), and applying conventional random forests using guided bagging [21]. Synthesis of complex data, such as cardiac diffusion images (DTI), was also done. Synthesis of complex shapes, using spectral graph decompositions, is currently on-going work.

The modeling of shapes also includes novel representations based on the spectral decomposition of images[4] which are more robust to large deformations when comparing multiple patients.

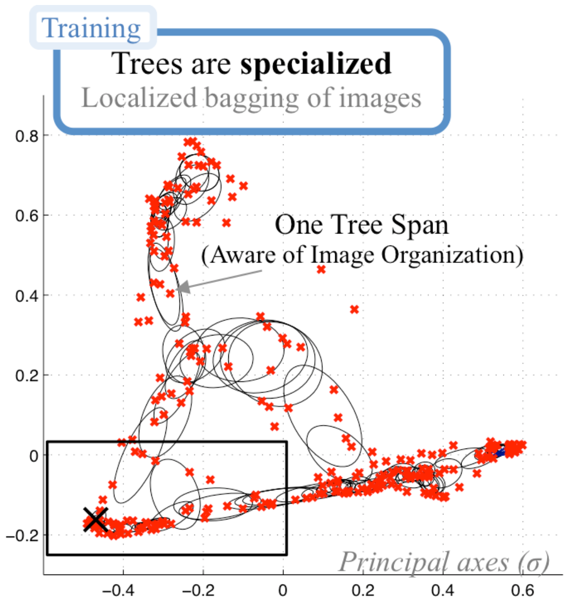


Figure 9. Laplacian Forest, where images are here represented as points, and where decision trees are trained using the spatial organization of these images on a reduced space.

5.3.3. Statistical analysis of heart shape, deformation and motion

Participants: Marc-Michel Rohé [correspondent], Xavier Pennec, Maxime Sermesant.

This work was partly supported by the FP7 European project MD-Paedigree and by ERC Advanced Grant MedYMA (on Biophysical Modeling and Analysis of Dynamic Medical Images)

Statistical analysis, Registration, Reduced order models, Machine learning

The work aims at developping statistical tools to analyse cardiac shape, deformation, and motion. In particular, we are interested in developping reduced order models so that the variability within a population described by a complex model can be reduced into few parameters or modes that are clinically relevant. We use these modes to represent the variability seen in a population and to relate this variability with clinical parameters, and we build group-wise statistics which relate these modes to a given pathology. We focus on cardiomyopathies and the cardiovascular disease risk in obese children and adolescents.

5.3.4. Geometric statistics for Computational Anatomy

Participants: Nina Miolane [Correspondent], Xavier Pennec.

Lie groups, pseudo-Riemannian, Statistics, Computational Anatomy

Lie groups are widely used in mathematical models for Medical Imaging. In Computational Anatomy for example, an organ's shape can be modeled as the deformation of a reference shape, in other words : as an element of a Lie group. If one wants to analyze the variability of the human anatomy, e.g. to help diagnose diseases, one has to perform statistics on Lie groups. We investigate the geometric structures on Lie groups that enable to define consistent statistics. A Lie group G is a manifold with an additional group structure. Statistics on *Riemannian* manifolds have be studied throughout the past years. One may wonder if we could use the theory of statistics on Riemannian manifolds for statistics on G. To this aim, we need to define a Riemannian metric on the Lie group that is *compatible with the group structure*: a so-called *bi-invariant* metric. However, it is known that most Lie groups do not admit any bi-invariant metric. One may wonder if we could generalize the theory of statistics on Riemannian manifolds to pseudo-Riemannian manifolds and use it for statistics on G. To this aim, we need to define a bi-invariant pseudo-metric on G. How many Lie groups do admit such a pseudo-metric and can we compute it? These investigations and their results (see Fig. 10) were presented at MaxEnt 2014 [24].

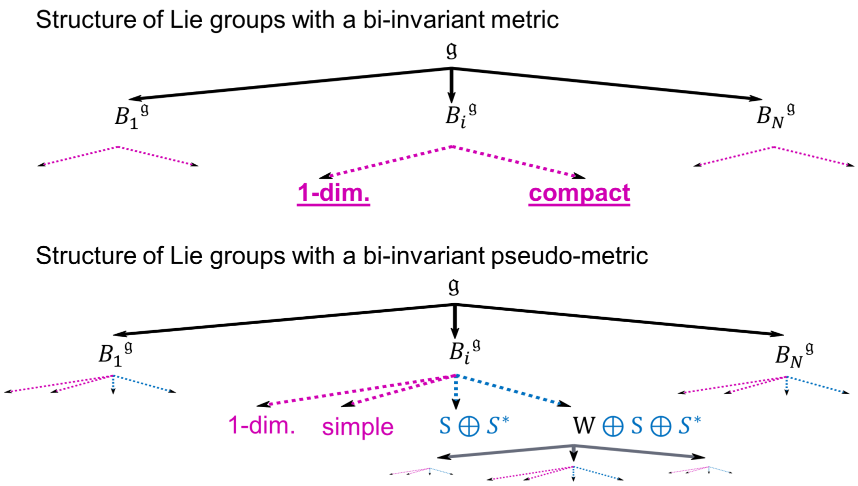


Figure 10. Structure of Lie groups on which one can define a bi-invariant metric or a bi-invariant pseudo-metric. The black levels of the tree represent the adjoint decomposition of the Lie algebra, the dashed lines represent the possible algebraic types of the substructures. Note the recursive construction in the pseudo-Riemannian case.

5.3.5. Statistical Analysis of Diffusion Tensor Images of the Brain

Participants: Vikash Gupta [correspondent], Nicholas Ayache, Xavier Pennec.

Population specific multimodal brain atlas for statistical analysis of white matter tracts on clinical DTI.

HIV virus can cross the hematoencephallic barrier and affect the neural connectivity in the human brain causing compromised motor controls, loss in episodic, long term memory and working memory, loss in attention/concentration and visual agnosia. These cognitive losses are characterized by the neuropsychological (NP) test scores and believed to be correlated with destruction of white matter (WM) integrity among the HIV patients. For quantifying the loss in WM integrity, the HIV subjects are compared against controls using a tract based spatial statistics (TBSS) routine. The standard TBSS routines uses univariate statistics using the fractional anisotropy (FA) maps. However, we improved on the existing routines using tensor based registration for normalizing the diffusion tensor images (DTI) followed by a multivariate statistics using the full tensor information. With the improved method it is possible to detect differences in WM regions which was not possible using the existing TBSS routines. For this study a population specific multimodal (T1 and DTI) brain atlas was developed from the population. The joint atlas also contains a probabilistic parcellation of WM regions in the brain which can be used for region of interest (ROI) based statistical studies (see Fig.11).

5.3.6. Longitudinal Analysis and Modeling of Brain Development

Participants: Mehdi Hadj Hamou [correspondent], Xavier Pennec, Nicholas Ayache.

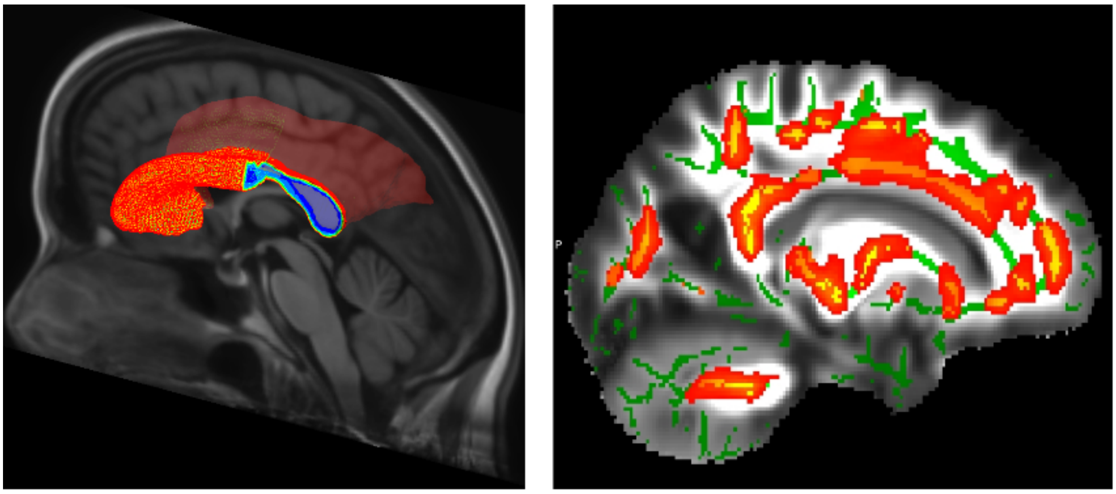


*Figure 11. **A:** Probabilistic parcellation of corpus callosum with blue and red being the maximum and minimum probability regions respectively. **B:** Multivariate statistics on white matter tracts. The red-yellow sections show statistically significant differences*

This work is partly funded through the ERC Advanced Grant MedYMA 2011-291080 (on Biophysical Modeling and Analysis of Dynamic Medical Images).

Brain development, adolescence, longitudinal analysis, non-rigid registration algorithm, extrapolation, interpolation

This work is divided into 2 complementary studies about longitudinal trajectories modeling:

- Diffeomorphic registration parametrized by Stationary Velocity Fields (SVF) is a promising tool already applied to model longitudinal changes in Alzheimer's disease. However, the validity of these model assumptions in faithfully describing the observed anatomical evolution needs to be further investigated. In this work, we thus analyzed the effectiveness of linear regression of SVFs in describing anatomical deformations estimated from past and future observations of the MRIs.
- Due to the lack of tools to capture the subtle changes in the brain, little is known about its development during adolescence. The aim of this project is to provide quantification and models of brain development during adolescence based on diffeomorphic registration parametrized by SVFs (see Fig.12). We particularly focused our study on the link between gender and the longitudinal evolution of the brain. This work was done in collaboration with J.L. Martinot et H. Lemaître (Inserm U1000).

5.4. Computational Physiology

5.4.1. Biophysical Modeling and Simulation of Longitudinal Brain MRIs with Atrophy in Alzheimer's Disease

Participants: Bishesh Khanal [correspondent], Nicholas Ayache, Xavier Pennec.

This work has been partly supported by the European Research Council through the ERC Advanced Grant MedYMA (on Biophysical Modeling and Analysis of Dynamic Medical Images).

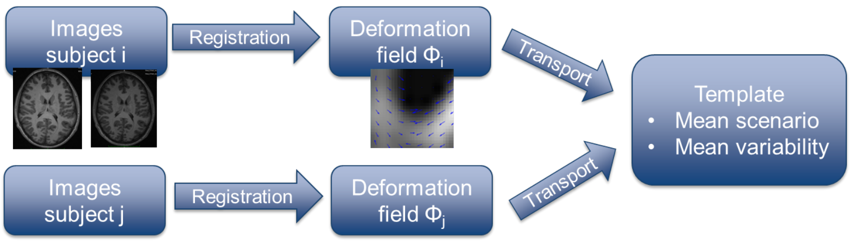


Figure 12. Pipeline for the longitudinal analysis of brain development during adolescence.

Alzheimer's Disease (AD), modeling brain deformation, biophysical model, simulation

- We propose a biophysical model of brain deformation due to atrophy in Alzheimer's Disease(AD) [17]. The model allows simulation of longitudinal brain MRIs with a desired level of atrophy in brain parenchyma. Here we enhanced our previous implementation to model brain parenchyma and cerebrospinal fluid (CSF) differently so that there is no need to prescribe atrophy in CSF region (see Fig.13).
- The model could be used to explore different possible hypotheses about evolution of atrophy in the brain and how it affects the brain shape changes.

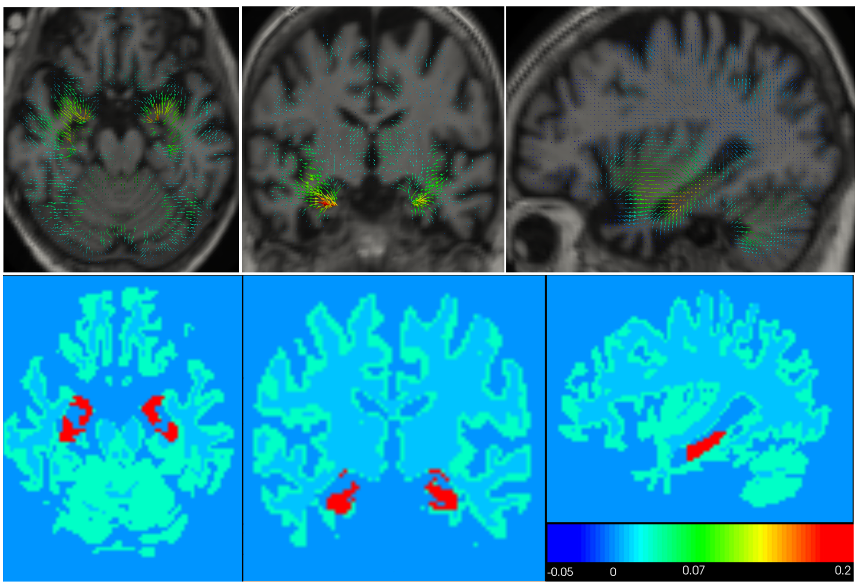


Figure 13. An example of obtained deformation field (top) from the model for the prescribed atrophy (bottom). From left to right: Axial, Coronal and Sagittal views.

5.4.2. Glioblastoma : Study of the vasogenic edema

Participants: Matthieu Lê [correspondent], Hervé Delingette, Jan Unkelbach [Massachussetts General Hospital], Nicholas Ayache.

This work is carried out between Ascelpios research group, Inria Sophia Antipolis, France and the Department of Radiation Oncology of the Massachusetts General Hospital, Boston, USA.

Glioblastoma, Vasogenic Edema, Radiotherapy, Target Delineation

- We studied the impact of anti-angiogenic treatment on the MRI appearance of glioblastoma.
- We studied how MRI extracted features could help distinguish between the vasogenic edema and the tumor infiltration[22].
- We analyzed the impact of excluding the vasogenic edema from the gross tumor volume during radiation therapy (see Fig. 14). Our approach leads to a dose more comformal to the underlying tumor cell density knowing that prescribing less dose might open the way for later re-irradiation.

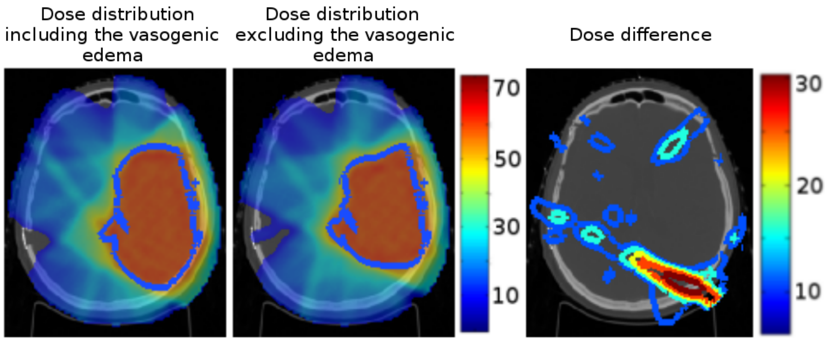


Figure 14. Comparison of the dose distribution including the vasogenic edema (clinical practice) and the excluding the estimated vasogenic edema (proposed method).

5.4.3. Image-based Prediction of Cardiac Ablation Targets

Participants: Rocio Cabrera Lozoya [correspondent], Maxime Sermesant, Nicholas Ayache.

Electrophysiology, ablation planning, machine learning

Ventricular radio-frequency ablation (RFA) can have a critical impact on preventing sudden cardiac arrest but is challenging due to a highly complex arrhythmogenic substrate. We aim at identifying local image characteristics capable of predicting the presence of local abnormal ventricular activities (LAVA). This could lead to pre-operatively and non-invasively improve and accelerate the procedure.

- We present the use of intensity and texture-based local imaging features in the vicinity of myocardial scar and grey zones towards the prediction of RFA target localisation (see Fig.15).
- We detail the uncertainty in the data and explore its impact on the classification results.
- A preliminary output with visual interpretation and potential use in a clinical environment was presented.
- The encouraging obtained results warrant further investigation and open up possibilities for non-invasive cardiac arrhythmia ablation planning. [13]

5.4.4. Personalised Canine Electromechanical Model of the Heart

Participants: Sophie Giffard-Roisin [correspondent], Maxime Sermesant, Hervé Delingette, Stéphanie Marchesseau, Nicholas Ayache.

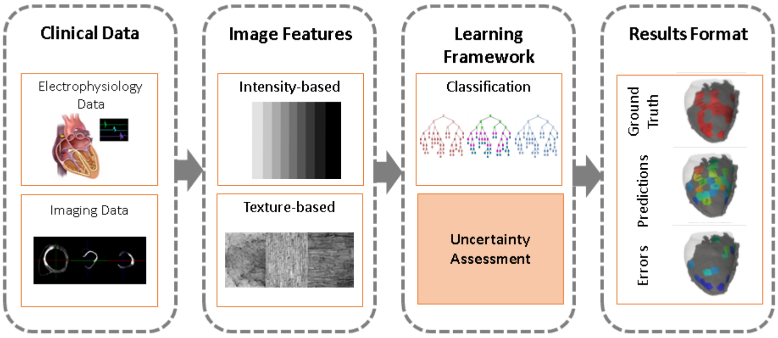


Figure 15. Pipeline showing the processing of our multimodal data. It includes an image feature extraction phase, followed by a classification with uncertainty assessment stage. The rightmost panel shows the preliminary output result format for a clinical environment.

This work has been supported by the European Project FP7 under grant agreement VP2HF (no 611823) and the ERC Advanced Grant MedYMA (on Biophysical Modeling and Analysis of Dynamic Medical Images).

Cardiac Modelling, Personalised Simulation, Electrical and Mechanical Simulation

- We studied the coupled electro-mechanical modelling of the heart, where the mechanics is handled by the Bestel-Clement-Sorine model while the electrophysiological phenomena is driven by an Eikonal model (see Fig. 16).
- We participated to the STACOM'2014 LV Mechanics Challenge in Boston[15] where four healthy canine clinical data (left ventricles) were provided. The validation was performed on local displacements. Our model has been calibrated by a quantitative sensitivity study as well as a personalized automatic calibration.

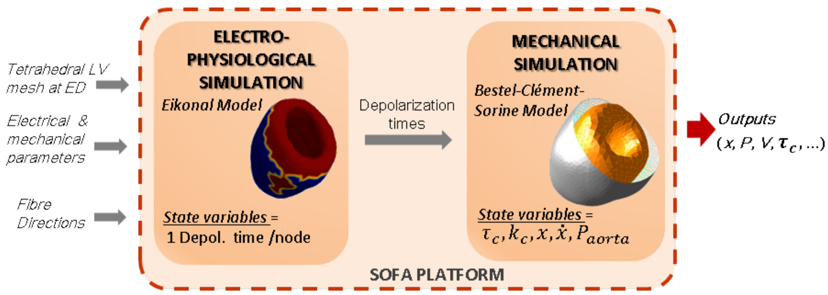


Figure 16. Complete electromechanical pipeline used for the simulations of four healthy canine hearts.

5.4.5. Computational modeling of radiofrequency ablation for the planning and guidance of abdominal tumor treatment

Participants: Chloé Audigier [correspondent], Hervé Delingette, Tommaso Mansi, Nicholas Ayache.

This PhD is carried out between Asclepios research group, Inria Sophia Antipolis and the Image Analytics and Informatics global field, Siemens Corporate Research, Princeton, USA.

Radio Frequency Abation, Patient-Specific Simulation, Lattice Boltzmann Method, Computational Fluid Dynamics, Heat Transfer, Therapy Planning, Liver

Radiofrequency abation (RFA) is a minimally invasive therapy suited for liver tumor ablation. However a patient-specific predicitive tool to plan and guide the treatment is required. We developed a computational framework for patient-specific planning of RFA (see Fig.17) :

- a personalised forward model of RFA:
- A patient-specific detailed anatomical model of the liver is estimated from standard CT image and meshed to generate a tetrahedral volume mesh. The structures of interest include the parenchyma, lesion, hepatic vein and vena cava.
- A Computation Fluic Dynamic and porous media solver using the Lattice Boltzmann Method is used to compute the patient-specific blood flow in the hepatic circulatory system and the blood flow distribution inside the parenchyma.
- Bio-heat equation has been implemented with a Lattice Boltmann Method also to model efficiently the heat propagation in biological tissues accounting for the cooling effect of neighboring vessels. A cell death model have been combined to account for the cellular necrosis.

Then this forward model is used to estimate patient-specific model parameters as presented in the ABDI workshop at MICCAI 2014 [12]. This work presented obtained the best paper award of the workshop.

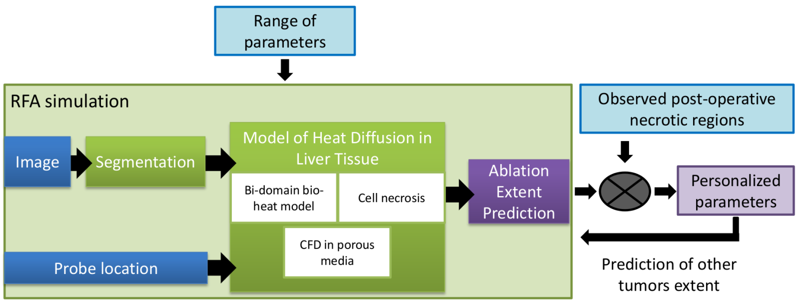


Figure 17. Steps of the proposed method for parameters personnalisation (blue: input, green: processes, purple: output).

5.4.6. Multi-channel patch-based glioma segmentation

Participants: Nicolas Cordier [correspondent], Hervé Delingette, Nicholas Ayache.

Part of this work was funded by the European Research Council through the ERC Advanced Grant MedYMA (on Biophysical Modeling and Analysis of Dynamic Medical Images).

Brain, MRI, Glioma, Patch-based Segmentation, Tumor Simulation

The segmentation of glioblastoma, the most severe case of brain tumors, is a crucial step for diagnostic assessment and therapy planning. In order to perform the manual delineation of the tumor compartments, the clinicians have to concurrently screen multi-channel 3D MRI, which makes the process both time-consuming and subject to inter-expert delineation variability. We are building upon the patch-based segmentation framework, the state-of-the-art for the segmentation of healthy brain structures, to present automatic glioma segmentation algorithms. Our 2013 submission to the MICCAI Brain Tumor Segmentation Challenge has been improved by:

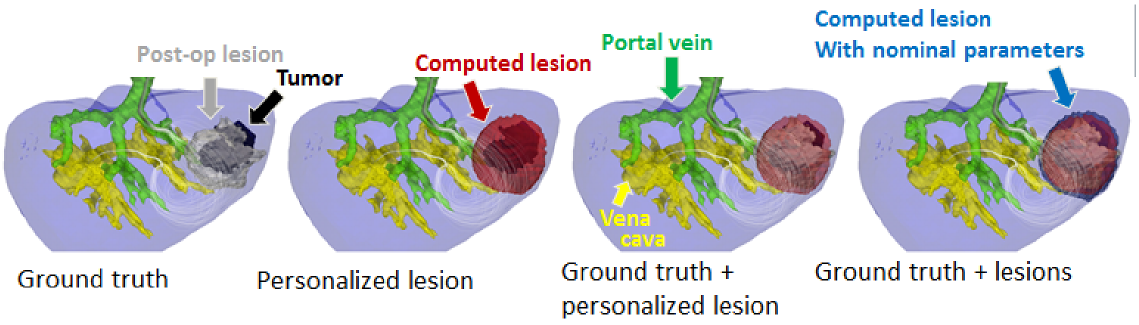


Figure 18. Predicted necrosis compared qualitatively well with ground truth (necrosis zone observed on a post-operative image).

- replacing the heuristic label fusion strategy with a more robust approach,
- integrating information such as statistics of appearance and position,
- generating configurations of synthetic training patches,
- filtering out the training patches for which the labels are less reliable (see Fig.19).

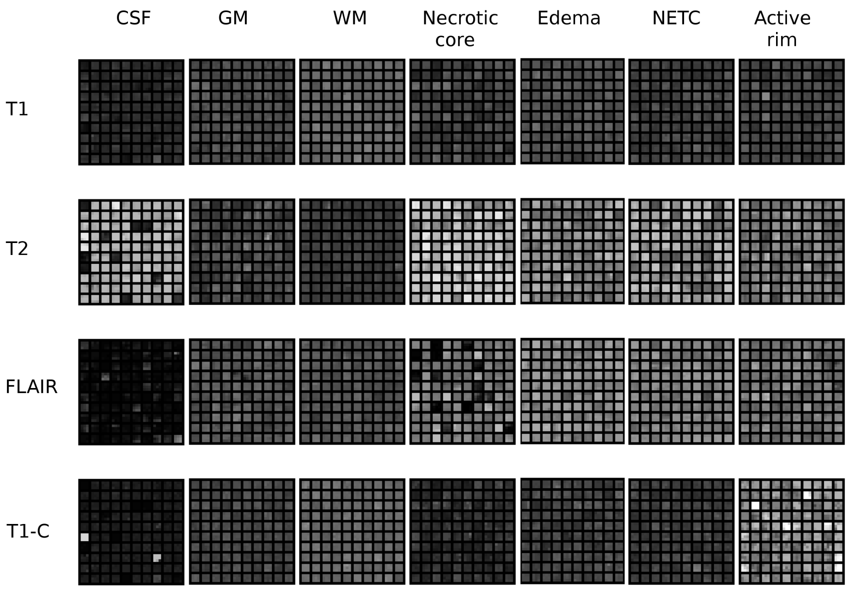


Figure 19. 2D slices of randomly sampled 3D multi-channel patches. From left to right: the different MR channels (T1, T2, T2-FLAIR, contrast enhanced T1). From top to bottom: cerbrospinal fluid (CSF), grey matter (GM), white matter (WM), necrotic tumor core, edema, non-enhancing tumor core (NETC), and active rim.

ASCOLA Project-Team

6. New Results

6.1. Highlights of the Year

Nicolas Tabareau was awarded a starting grant from the European Research Council (ERC), the most prestigious type of research projects of the European Union for young researchers. From 2015–2020 he will pursue research on “CoqHoTT: Coq for Homotopy Type Theory.”

Jonathan Pastor has won the joint 1st prize at the Grid5000 Scale challenge, an international challenge for large-scale experiments on geographically-distributed cluster environments. Jonathan has shown with a colleague how to deploy and manage thousands of VMs in such an environment using his approach to fully distributed virtual machine management.

This year we have provided major research results in two domains. First, we have developed several new approaches for the formal reasoning over software in the domains of theorem proving [31], as well as reasoning over distributed interaction protocols [32] and software compositions [24]. Second, we have developed new methods supporting dynamic computations over the cloud, both by means of more elastic cloud applications [27] and better locality management for the dynamic placement of virtual machines in Cloud infrastructures [29].

6.2. Programming Languages

Participants: Ronan-Alexandre Cherrueau, Rémi Douence, Hervé Grall, Thomas Ledoux, Florent Marchand de Kerchove de Denterghem, Jacques Noyé, Jean-Claude Royer, Mario Südholt.

6.2.1. Formal Methods, logics and type theory

This year we have published new results extending previous type theories: we have introduced a notion of universe polymorphism for the theorem prover Coq and new type-based mechanisms for the definition and analysis of program equivalences. We have also shown how to harness capabilities, well-known in the security domain, in the context of the functional programming language Haskell. These results are detailed in the current section.

Furthermore, we have applied formal methods and typing in the context of aspect oriented programming ([12], [16], [24]) and in the context of distributed programming (aspectual session types [32]). We have also developed a framework for the formal definition and analysis of accountability properties based on temporal logics. These different results are detailed in Sec. 6.3 for details.

6.2.1.1. Universe Polymorphism in Coq

Universes are used in type theory to ensure consistency by checking that definitions are well-stratified according to a certain hierarchy. In the case of the Coq proof assistant, based on the predicative Calculus of Inductive Constructions (pCIC), this hierarchy is built from an impredicative sort Prop and an infinite number of predicative Type universes. A cumulativity relation represents the inclusion order of universes in the core theory. Originally, universes were thought to be floating levels, and definitions to implicitly constrain these levels in a consistent manner. This works well for most theories, however the globality of levels and constraints precludes generic constructions on universes that could work at different levels. We have introduced universe polymorphism [31] that extends this setup by adding local bindings of universes and constraints, supporting generic definitions over universes, reusable at different levels. This provides the same kind of code reuse facilities as ML-style parametric polymorphism. However, the structure and hierarchy of universes is more complex than bare polymorphic type variables.

6.2.1.2. A Logical Study of Program Equivalence

Proving program equivalence for a functional language with references is a notoriously difficult problem. The goal of the thesis of Guilhem Jaber on "A Logical Study of Program Equivalence" [G. Jaber, Mines Nantes, July 14] was to propose a logical system in which such proofs can be formalized, and in some cases inferred automatically. In the first part, a generic extension method of dependent type theory has been proposed, based on a forcing interpretation seen as a presheaf translation of type theory. This extension equips type theory with guarded recursive constructions, which are subsequently used to reason on higher-order references. In the second part, he has defined a nominal game semantics for a language with higher-order references. It marries the categorical structure of game semantics with a trace representation of denotations of programs, which can be computed operationally and thus have good modularity properties. Using this semantics, he has proven completeness of Kripke logical relations defined in a direct way, using guarded recursive types, without using biorthogonality. The problem of contextual equivalence is then reduced to the satisfiability of an automatically generated formula defined in this logic, that is, to the existence of a world validating this formula. Under some conditions, this satisfiability can be decided using a SMT solver.

6.2.1.3. Effect Capabilities For Haskell

Computational effects complicate the tasks of reasoning about and maintaining software, due to the many kinds of interferences that can occur. While different proposals have been formulated to alleviate the fragility and burden of dealing with specific effects, such as state or exceptions, there is no prevalent robust mechanism that addresses the general interference issue. Building upon the idea of capability-based security, we have proposed effect capabilities [25] as an effective and flexible manner to control monadic effects and their interferences. Capabilities can be selectively shared between modules to establish secure effect-centric coordination. We have further refined capabilities with type-based permission lattices to allow fine-grained decomposition of authority. An implementation of effect capabilities in Haskell has been done, using type classes to establish a way to statically share capabilities between modules, as well as to check proper access permissions to effects at compile time.

6.2.2. Language Mechanisms

In 2014, we have proposed new general language-based mechanisms for concurrent event-based systems and sequential programming languages. Moreover, we have investigated domain-specific languages that support aspect-oriented programming and provide control over propagation strategies in constraint solvers. These results are detailed in the remainder of this section.

Furthermore, we have proposed language support for the definition and enforcement of security properties, in particular related to the accountability of service-based systems, see Sec. 6.3 .

6.2.2.1. Concurrent Event-Based Programming

Advanced concurrency abstractions overcome the drawbacks of low-level techniques such as locks and monitors, freeing programmers that implement concurrent applications from the burden of concentrating on low-level details. However, with current approaches the coordination logic involved in complex coordination schemas is fragmented into several pieces including join patterns, data emissions triggered in different places of the application, and the application logic that implicitly creates dependencies among communication channels, hence indirectly among join patterns. In [33], we have presented JEScala, a language that captures coordination schemas in a more expressive and modular way by leveraging a seamless integration of an advanced event system with join abstractions. We have validated the approach with case studies and provided a first performance assessment.

6.2.2.2. Lazy imperative programming

Laziness is a powerful concept in functional programming that permits the reuse of general functions in a specific context, while keeping performance close to the efficiency of dedicated definitions. Lazy evaluation can be used in imperative programming too. Twenty years ago, John Launchbury was already advocating for lazy imperative programming, but the level of laziness of his framework remained limited. Twenty years after, the picture has not changed.

We have proposed an Haskell framework to specify computational effects of imperative programs as well as their dependencies [23]. We have presented a semantics of a call-by-need lambda-calculus extended with imperative strict and lazy features and proved the correctness of our approach. While originally motivated by a less rigid use of foreign functions, we have shown that our approach is fruitful for a simple scenario based on sorted mutable arrays. Furthermore, we can take advantage of equations between algebraic operations to dynamically optimize compositions of imperative computations.

6.2.2.3. Domain-Specific Aspect Languages

Domain-Specific Aspect Languages (DSALs) are Domain-Specific Languages (DSLs) designed to express crosscutting concerns. Compared to DSLs, their aspectual nature greatly amplifies the language design space. In the context of the Associate Team RAPIDS/REAL, we have structured this space in order to shed light on and compare the different domain-specific approaches to deal with crosscutting concerns [37]. We have reported on a corpus of 36 DSALs covering the space, discussed a set of design considerations and provided a taxonomy of DSAL implementation approaches. This work serves as a frame of reference to DSAL and DSL researchers, enabling further advances in the field, and to developers as a guide for DSAL implementations.

6.2.2.4. Controlling constraint propagation

Constraint propagation is at the heart of constraint solvers. Two main trends co-exist for its implementation: variable-oriented propagation engines and constraint-oriented propagation engines. These two approaches ensure the same level of local consistency but their efficiency (computation time) can be quite different depending on the problem instances to be solved. However, it is usually accepted that there is no best approach in general, and modern constraint solvers implement only one of them.

In the context of Charles Prud'homme's PhD Thesis [15], we have gone a step further providing a solver independent language at the modeling stage to enable the design of propagation engines. We have validated our proposal with a reference implementation based on the Choco solver and the MiniZinc constraint modeling language.

6.3. Software Composition

Participants: Diana Allam, Walid Benghabrit, Ronan-Alexandre Cherrueau, Rémi Douence, Hervé Grall, Thomas Ledoux, Jean-Claude Royer, Mohamed Sellami, Mario Südholt.

6.3.1. Constructive Security

Nowadays we are witnessing the wide-spread use of cloud services. As a result, more and more end-users (individuals and businesses) are using these services for achieving their electronic transactions (shopping, administrative procedures, B2B transactions, etc.). In such scenarios, personal data is generally flowing between several entities and end-users need (i) to be aware of the management, processing, storage and retention of personal data, and (ii) to have necessary means to hold service providers accountable for the usage of their data. Usual preventive security mechanisms are not adequate in a world where personal data can be exchanged on-line between different parties and/or stored at multiple jurisdictions. Accountability becomes a necessary principle for the trustworthiness of open computer systems. It regards the responsibility and liability for the data handling performed by a computer system on behalf of an organization. In case of misconduct (e.g. security breaches, personal data leak, etc.), accountability should imply remediation and redress actions, as in the real life.

In 2014, we have developed two general approaches for the definition and enforcement of accountability properties.

6.3.1.1. Logic-based accountability properties

We have proposed a framework for the representation of cloud accountability policies [19]. Such policies offer end-users a clear view of the privacy and accountability obligations asserted by the entities they interact with, as well as means to represent their preferences. This framework comes with two novel accountability policy languages; an abstract one, which is devoted for the representation of preferences/obligations in an human readable fashion, a concrete one for the mapping to concrete enforceable policies. We motivate our solution with concrete use case scenarios. [30] discusses issues related to data privacy and big data technologies and advocate the use of the framework to support accountability.

We have provided an abstract language for the representation of accountability obligations [20]. We define its semantics using first-order temporal logic and a specific modality for accountability is introduced. We analyze a healthcare use case to illustrate the efficiency of our approach in representing accountability obligations in realistic situations. The use of such services-based applications usually implies the flow of personal data on-line between several parties. In [21], we consider this issue at the design-time of the software and we propose some foundations for an accountable software design. Accountability for a software is a property describing, among other aspects, its liability to end-users for the usage of the data it has been entrusted. We propose to enrich software's component design by accountability clauses using an abstract accountability language (introduced in [20]). We also define conditions for the well-formedness of an accountable component design and show how they can be checked using the μ-CRL model-checker.

6.3.1.2. Defining and enforcing multi-level accountability properties

Many accountability policies require access to all levels of the software stack of service-based applications. Furthermore, they should include explicit means for the definition of cross-domain policies and provide constructive means for the implementation of a wide variety of of accountability properties. These features, in particular, multi-level support, are missing in existing approaches.

We have provided an approach that addresses these objectives explicitly through a language for the definition of expressive regular policies over accountability predicates applicable at all levels of the service stack [22]. Furthermore, we have presented hierarchies of constructive schemes for the implementation of policies for transparency and remediation properties that are implemented in terms of our accountability policy language. Finally, we have shown how to harness the accountability schemes to tackle real-world violations of accountability properties arising from security vulnerabilities of OAuth-based authorization and authentication protocols.

6.3.2. Aspect-Oriented Programming

We have produced in 2014 a range of results enabling reasoning over aspect languages and investigated the use of execution levels. These results are presented in the remainder of this section.

We have also applied ideas from aspect oriented programming in the context of distributed programming (aspectual session types [32]), see Sec. 6.4 .

6.3.2.1. Reasoning about aspect interference using effective aspects

Aspect-oriented programming (AOP) aims at enhancing modularity and reusability in software systems by offering an abstraction mechanism to deal with crosscutting concerns. But, in most general-purpose aspect languages aspects have almost unrestricted power, eventually conflicting with these goals. To tame aspects, we have proposed Effective Aspects: a novel approach to embed the pointcut/advice model of AOP in a statically-typed functional programming language like Haskell; along two main contributions. First, we have defined a monadic embedding of the full pointcut/advice model of AOP [16].

Type soundness is guaranteed by exploiting the underlying type system, in particular phantom types and a new anti-unification type class. In this model aspects are first-class, can be deployed dynamically, and the pointcut language is extensible, therefore combining the flexibility of dynamically-typed aspect languages with the guarantees of a static type system. Monads (which allow the definition of sequences of computations in functional programs) enable us to directly reason about computational effects both in aspects and base programs using traditional monadic techniques. Using this we extend the notion of Open Modules with effects, and also with protected pointcut interfaces to external advising. These restrictions are enforced statically using the type system. Also, we adapt the techniques of EffectiveAdvice to reason about and enforce control flow properties as well as to control effect interference. We show that the parametricity-based approach to effect interference falls short in the presence of multiple aspects and propose a different approach using monad views, a novel technique for handling the monad stack, developed by Schrijvers and Oliveira. Then, we exploit the properties of our model to enable the modular construction of new semantics for aspect scoping and weaving. Our second contribution [24] builds upon a powerful model to reason about mixin-based composition of effectful components and their interference, based on equational reasoning, parametricity, and algebraic laws about monadic effects. Our contribution is to show how to reason about interference in the presence of unrestricted quantification through pointcuts. We show that global reasoning can be compositional, which is key for the scalability of the approach in the face of large and evolving systems. A comprehensive version of those two works appears in Ismael Figueroa PhD thesis [12].

6.3.2.2. Execution Levels for AOP: from program design to applications

In AOP languages, advice evaluation is usually considered as part of the base program evaluation. This is also the case for certain pointcuts, such as if pointcuts in AspectJ, or simply all pointcuts in higher-order aspect languages like AspectScheme. While viewing aspects as part of base level computation clearly distinguishes AOP from reflection, it also comes at a price: because aspects observe base level computation, evaluating pointcuts and advice at the base level can trigger infinite regression. To avoid these pitfalls, aspect languages propose ad hoc mechanisms, which increase the complexity for programmers while being insufficient in many cases. We have proposed to clarify the situation by introducing levels of execution in the programming language [18], thereby allowing aspects to observe and run at specific, possibly different, levels. We have adopted a defensive default that avoids infinite regression, and gives advanced programmers the means to override this default using level-shifting operators.

6.3.3. Service provisioning

This year, we have provided results on two fundamental problems of service-oriented architectures: service interoperability and service mediation.

6.3.3.1. Service interoperability

Web service support a document-oriented style for clients to interact with a server and promote an environment for systems that is loosely coupled and interoperable. Two models exist for implementing Web services: A process-oriented Web services model, SOAP, and a resource-oriented Web services model, RESTful. Service components are mainly based on description interfaces. These interfaces are often known as structural standardized interfaces like WSDL for SOAP and WADL for RESTful. The implementation of Web services is increasingly based on object-oriented (OO) frameworks, at the client and the server sides. Using these frameworks, developers can transform an object code into a Web service, or access a remote Web service, at the touch of a button. In this context, two levels are present: an object level built over a service level.

Diana Allam's PhD thesis [11] has focused on two properties of these frameworks:

- The loose coupling between the two levels, which allows the complex technical details of the service level to be hidden at the object level and the service level to be evolved with a minimal impact on the object level.
- The interoperability induced by the substitution principle associated to subtyping in the object level, which allows to freely convert a value of a subtype into a value of a supertype.

The thesis provides three contributions in this context. We propose a unified formal model for web services based on message passing and enabling first class channels. It is equipped with a powerful type-checking allowing union, intersection and negation operations as well as subtyping. The type checking algorithm relies on the semantic approach defined by G. Castagna. This type system is also protected against attackers. The second contribution is a concrete refinement of the model into RESTful and SOAP frameworks as well as a unified API for service discovery. To define such an API, we have first shown how the details of the standard interfaces (WSDL and WADL) could be simplified and abstracted and then we rely on subtyping in the discovery mechanism. Finally, to solve some of the interoperability issues between the OO level and the service level a formalization of the binding using categorical concepts (commutative diagrams) is proposed. Based on this an analysis of the mismatch problems has been done and a new specification of the data binding has been formalized. The document then discusses some variations in the implementation of the data binding solution and a prototype for the Apache CXF framework.

Mayleen Lacouture's PhD thesis "A Chemical Programming Language for Orchestrating Services - Application to Interoperability Problems" [M. Lacouture, MN/U. Nantes, Oct. 14] proposes a framework easing interoperability in the form of an architecture that integrates different orchestration languages with heterogeneous service providers around a pivot language. The pivot language is implemented as a new orchestration language based on the chemical programming paradigm. Concretely, the dissertation presents a language called Criojo that implements and extends the Heta-calculus, an original calculus associated to a chemical abstract machine dedicated to service-oriented computing. The consequence of adopting this approach would be an improvement in the interoperability of services and orchestration languages, thus easing the development of composite services. The high level of abstraction of Criojo could allow developers to write very concise orchestrations since message exchanges are represented in a natural and intuitive way.

6.3.3.2. Service mediation

Service composition is a major advance service-oriented computing brings to enable the development of distributed applications. However, the distributed nature of services hampers their composition with data heterogeneity problems. We address these problems with a decentralized Mediation-as-a-Service architecture that solves data inconsistencies occurring during the composition of business services [17]. As an extension to our previous work that focused on data interpretation problems, we present in this paper a solution to solve data inconsistencies at the syntactic, structural and semantic levels. We show how syntactic, structural and semantic mediation techniques can be combined, and how semantic mediation provides useful information that helps structural and syntactic mediation. We demonstrate how our architecture enables decentralized publication and discovery of mediation services. We motivate our work with a concrete scenario and validate our proposal with experiments.

6.3.4. Software product line architectures

Software product lines were designed from the product line tested out by H. Ford at the beginning of the 20TH century, which led to the success of his automotive production. For 15 years, these methods have been visible in several software application fields: telephony at Nokia, televisions at Philips, print software at HP and flight applications at Boeing, among others. The concept of architecture is crucial for classic software applications, and this concept is even more important at the level of domain engineering in product lines. In a product line, the so-called reference architecture generically describes the architectures of all the products in the family. The chapter [34] describes the technical means and methods for defining a reference architecture for a software product line. It also presents the methods for operating this architecture through, for example, techniques emerging from model and software component engineering, or aspect-oriented programming. These concepts and techniques are illustrated using a case study.

6.4. Cloud applications and infrastructures

Participants: Adrien Lebre, Thomas Ledoux, Yousri Kouki, Guillaume Le Louët, Jean-Marc Menaud, Jonathan Pastor, Flavien Quesnel, Mario Südholt.

In 2014, we have provided solutions for Cloud-based and distributed programming, virtual environments and data centers, in particular concerning energy-optimal Cloud applications.

6.4.1. Cloud and distributed programming

This year we have published results on a broker that provides better guarantees on service-level agreements in the Cloud. Furthermore, we have extended a class of formally-defined protocols, session types.

6.4.1.1. Service-level agreement for the Cloud

Elasticity is the intrinsic element that differentiates Cloud Computing from traditional computing paradigms, since it allows service providers to rapidly adjust their needs for resources to absorb the demand and hence guarantee a minimum level of Quality of Service (QoS) that respects the Service Level Agreements (SLAs) previously defined with their clients. However, due to non-negligible resource initiation time, network fluctuations or unpredictable workload, it becomes hard to guarantee QoS levels and SLA violations may occur.

We propose a language support for Cloud elasticity management that relies on CSLA (Cloud Service Level Agreement) [27]. CSLA offers new features such as QoS/functionality degradation and an advanced penalty model that allow providers to finely express contracts so that services self-adaptation capabilities are improved and SLA violations minimized. The approach was evaluated with a real infrastructure and application testbed. Experimental results show that the use of CSLA makes Cloud services capable of absorbing more peaks and oscillations by trading-off the QoS levels and costs due to penalties.

6.4.1.2. AO session types for distributed protocols

Multiparty session types allow the definition of distributed processes with strong communication safety properties. A global type is a choreographic specification of the interactions between peers, which is then projected locally in each peer. Well-typed processes behave accordingly to the global protocol specification. Multiparty session types are however monolithic entities that are not amenable to modular extensions. Also, session types impose conservative requirements to prevent any race condition, which prohibit the uniform application of extensions at different points in a protocol. We have proposed a means to support modular extensions with aspectual session types [32], a static pointcut/advice mechanism at the session type level. To support the modular definition of crosscutting concerns, we have augmented the expressivity of session types to allow harmless race conditions. As a result, aspectual session types make multiparty session types more flexible, modular, and extensible.

6.4.2. Virtualization and data centers

In 2014, we have produced a variety of results on a new model for utility computing that addresses fundamental shortcomings of today's Cloud computing model. Furthermore, we have provided more powerful techniques for the virtualization of computations and the management of cluster-based environments, such as data centers.

6.4.2.1. Next generation utility computing

To accommodate the ever-increasing demand for Utility Computing (UC) resources while taking into account both energy and economical issues, the current trend consists in building larger and larger data centers in a few strategic locations. Although such an approach enables to cope with the actual demand while continuing to operate UC resources through centralized software system, it is far from delivering sustainable and efficient UC infrastructures. Throughout the Discovery initiative [0], we investigate how UC resources can be managed differently, considering locality as a primary concern. Concretely, we study how it can be possible to leverage any facilities available through the Internet in order to deliver widely distributed UC platforms that can better match the geographical dispersal of users as well as the unending resource demand. Critical to the emergence of such locality-based UC (LUC) platforms is the availability of appropriate operating mechanisms. We presented a prospective vision of a unified system driving the use of resources at an unprecedented scale by turning a complex and diverse infra structure into a collection of abstracted computing facilities that is both easy to operate and reliable [35]. By deploying and using such a LUC Operating System on backbones, our ultimate vision is to make possible to host/operate a large part of the Internet by its internal structure itself: A scalable and nearly infinite set of resources delivered by any computing facilities forming the Internet, starting from the larger hubs operated by ISPs, governments and academic institutions to any idle resources that may be provided by end-users. We highlight that this work is conducted through a collaboration between the ASAP, ASCOLA, AVALON and MYRIADS Inria Project-teams.

6.4.2.2. Adding locality capabilities to virtual machine schedulers

Through the DVMS proposal, we showed in 2013 the benefit of leveraging peer-to-peer algorithms to design and implement virtual machines (VMs) scheduling algorithms. Although P2P based proposals considerably improve the scalability, leading to the management of hundreds of thousands of VMs over thousands of physical machines (PMs), they do not consider the network overhead introduced by multi-site infrastructures. This over- head can have a dramatic impact on the performance if there is no mechanism favoring intra-site v.s. inter-site manipulations. This year, we extended our DVMS mechanism with a new building block designed on top of the Vivaldi coordinates mechanism. We showed its benefits by discussing several experiments performed

[0] http://beyondtheclouds.github.io

on four distinct sites of the Grid'5000 testbed. With our proposal and without changing the scheduling decision algorithm, the number of inter-site operations has been reduced by 72% [29]. This result provides a glimpse of the promising future of using locality properties to improve the performance of massive distributed Cloud platforms. We highlight that this work has been performed in collaboration with the ASAP, ASCOLA, AVALON and MYRIADS Inria Project-teams.

6.4.2.3. WAN-wide elasticity capabilities for distributed file systems

Applications dealing with huge amounts of data suffer significant performance impacts when they are deployed on top of an hybrid platform (i.e the extension of a local infrastructure with external cloud resources). More precisely, through a set of preliminary experiments we shew that mechanisms which enable on demand extensions of current Distributed File Systems (DFSes) are required. These mechanisms should be able to leverage external storage resources while taking into account the performance constraints imposed by the physical network topology used to interconnect the different sites. To address such a challenge we presented the premises of the Group Based File System, a glue providing the elasticity capability for storage resources by federating on demand any POSIX file systems [28].

6.4.3. Energy optimization

Demand for Green services is increasing considerably as people are getting more environmental conscious to build a sustainable society. Therefore, enterprise and clients want to shift their workloads towards green Cloud environment offered by the Infrastructure-as-a-Service (IaaS) provider. The main challenge for an IaaS provider is to determine the best trade-off between its profit while using renewable energy and customers satisfaction. In order to address this issue, we propose a *Cloud energy broker* [26], which can adjust the availability and price combination to buy Green energy dynamically from the market to make datacenter green. Our energy broker tries to maximize of using renewable energy under strict budget constraint whereas it also tries to minimize the use of brown energy by capping the limit of overall energy consumption of datacenter. The energy broker was evaluated with a real workload traced by PlanetLab. Experimental results show that our energy broker successfully enables meeting the best trade-off.

ASPI Project-Team

5. New Results

5.1. Adaptive multilevel splitting

Participants: Frédéric Cérou, Arnaud Guyader.

We show [21] that an adaptive version of multilevel splitting for rare events is strongly consistent. We also show that the estimates satisfy a CLT (central limit theorem), with the same asymptotic variance as the non-adaptive algorithm with the optimal choice of the parameters. It is a strong and general result, that generalizes some of our previous results, and the proof is quite technical and involved.

This work has been presented at the 10th International Workshop on Rare Event Simulation (RESIM), held in Amsterdam in August 2014.

5.2. Convergence of a two–step multilevel splitting algorithm for rare event simulation

Participants: François Le Gland, Damien–Barthélémy Jacquemart.

The problem is to accurately estimate the (very small) probability that a rare but critical event (such as a score function exceeding a given threshold) occurs before some fixed final time. Multilevel splitting is a very efficient solution, in which sample paths are propagated and are eliminated or replicated when some intermediate events (defined by some intermediate thresholds) occur. A common and efficient design is to define the next intermediate level as an empirical quantile of the running maximum of the score function along a surviving trajectory. However, it is practically impossible to remember when (at which time instant) and where (in which state) did each successful trajectory cross the empirically defined threshold. The proposed design is a two–step adaptive multilevel splitting algorithm: In the first step, a first set of trajectories is sampled in order to obtain the next intermediate threshold as an empirical quantile. In the second step, once the new intermediate threshold is obtained, a second set of trajectories is sampled in order to evaluate the transition probability to the new empirically defined intermediate region. This two–step procedure is repeated until some trajectories do hit the critical region before final time.

This work has been presented at the 10th International Workshop on Rare Event Simulation (RESIM), held in Amsterdam in August 2014.

5.3. Simulation–based algorithms for the optimization of sensor deployment

Participant: François Le Gland.

This is a collaboration with Christian Musso (ONERA, Palaiseau) and with Sébastien Paris (LSIS, université du Sud Toulon Var), related with the supervision of the PhD thesis of Yannick Kenné.

The problem considered here can be described as follows: a limited number of sensors should be deployed by a carrier in a given area, and should be activated at a limited number of time instants within a given time period, so as to maximize the probability of detecting a target (present in the given area during the given time period). There is an information dissymmetry in the problem: if the target is sufficiently close to a sensor position when it is activated, then the target can learn about the presence and exact position of the sensor, and can temporarily modify its trajectory so as to escape away before it is detected. This is referred to as the target intelligence. Two different simulation–based algorithms have been designed to solve separately or jointly this optimization problem, with different and complementary features. One is fast, and sequential: it proceeds by running a population of targets and by dropping and activating a new sensor (or re–activating a sensor already available) where and when this action seems appropriate. The other is slow, iterative, and non–sequential; it proceeds by updating a population of deployment plans with guaranteed and increasing criterion value at each iteration, and for each given deployment plan, there is a population of targets running to evaluate the criterion. Finally, the two algorithms can cooperate in many different ways, to try and get the best of both approaches. A simple and efficient way is to use the deployment plans provided by the sequential algorithm as the initial population for the iterative algorithm.

This work has been presented at the Conference on Optimization and Practices in Industry (COPI), held in Palaiseau in October 2014.

5.4. Non–homogeneous Markov switching auto–regressive models for wind time series

Participants: Valérie Monbet, Julie Bessac.

This is a collaboration with Pierre Ailliot (UBO) and Françoise Pène (UBO).

We proposed [20] non–homogeneous Markov switching auto–regressive models for bivariate wind time series considering Cartesian coordinates on one hand and polar coordinates on the other hand. In non–homogeneous models, the transitions depend on the wind direction at the previous time. At the location of interest, wind is rotating more often clockwise but wind direction may also oscillate around two prevailing directions (northeast for anti–cyclonic conditions and southwest for cyclonic conditions). These features induce respectively some cycles which can be seen in the second order structure and modes in the marginal distribution. In broad outline, non–homogeneous transitions help the process to stay in the same weather regime when the wind direction is close to the prevailing directions and lead to sojourn duration in the regimes which are not geometric.

5.5. Gaussian state–space models for wind speed

Participants: Valérie Monbet, Julie Bessac.

This is a collaboration with Pierre Ailliot (UBO).

A multi–site stochastic generator for wind speed has been developped [11]. It aims at simulating realistic wind conditions with a focus on reproducing the space-time motions of the meteorological systems. A Gaussian linear state–space model is used where the latent state may be interpreted as regional wind conditions and the observation equation links regional and local scales. The model is fitted to 6–hourly reanalysis data in the North–East Atlantic. It is shown that it is interpretable and provides a good description of important properties of the space–time covariance function of the data, such as the non full–symmetry induced by prevailing flows in this area.

5.6. Level–dependent time deformation of Gaussian processes

Participant: Valérie Monbet.

Many records in environmental science exhibit asymmetries. In this project, we introduce a time deformation to produce asymmetric path from a Gaussian process with symmetric path. A simple case is obtained by assuming that

$$Z_t = Y_{\phi(t)}, \quad \phi(t) = \int_0^t f(Z_s) ds$$

with $\{Y_t\}$ a stationary Gaussian process. The function f which controls the time deformation is increasing. The time–change function ϕ is such that the modified time increases quicker when the process is at high levels and thus that the crests of the modified process $\{Z_t\}$ are narrower than the ones of $\{Y_t\}$. The opposite holds true for the troughs. Inference tools are developed to estimate the function f.

5.7. Self–similar prior and wavelet bases for hidden turbulent motion

Participant: Patrick Héas.

This is a collaboration with Frédéric Lavancier (université de Nantes) and Souleymane Kadri–Harouna (université de la Rochelle)

This work [14] is concerned with the ill–posed inverse problem of estimating turbulent flows from the observation of an image sequence. From a Bayesian perspective, a divergence–free isotropic fractional Brownian motion (fBm) is chosen as a prior model for instantaneous turbulent velocity fields. This self–similar prior characterizes accurately second–order statistics of velocity fields in incompressible isotropic turbulence. Nevertheless, the associated maximum a posteriori involves a fractional Laplacian operator which is delicate to implement in practice. To deal with this issue, we propose to decompose the divergence–free fBm on well–chosen wavelet bases. As a first alternative, we propose to design wavelets as whitening filters. We show that these filters are fractional Laplacian wavelets composed with the Leray projector. As a second alternative, we use a divergence–free wavelet basis, which takes implicitly into account the incompressibility constraint arising from physics. Although the latter decomposition involves correlated wavelet coefficients, we are able to handle this dependence in practice. Based on these two wavelet decompositions, we finally provide effective and efficient algorithms to approach the maximum a posteriori.

5.8. Estimation of non–linear dynamics under sparse constraints

Participant: Patrick Héas.

This is a collaboration with Cédéric Herzet (EPI FLUMINANCE, Inria Rennes–Bretagne Atlantique) and Angélique Drémeau (ENSTA Bretagne, Brest).

Following recent contributions in non–linear sparse representations, this work [19], [18] focuses on a particular non–linear model, defined as the nested composition of functions. This family includes in particular discrete–time hidden Markov models. Recalling that most linear sparse representation algorithms can be straightforwardly extended to non–linear models, we emphasize that their performance highly relies on an efficient computation of the gradient of the objective function. In the particular case of interest, we propose to resort to a well–known technique from the theory of optimal control to evaluate the gradient. This computation is then implemented into the ℓ_1–reweighted procedure proposed by Candès et al. [24], leading to a non–linear extension of it. As an example, we consider the problem of estimating the ocean state from satellite low–dimensional information by exploiting a geophysical dynamical model and a sparse decomposition of the initial condition in some redundant dictionary.

This work has also been presented at Congrès National d'Assimilation, a national event held in Toulouse in December 2014.

ATEAMS Project-Team

5. New Results

5.1. Highlights of the Year

- Davy Landman, Jurgen Vinju received a Best paper award nomination, for their paper "Empirical analysis of the relationship between CC and SLOC in a large corpus of Java methods"(ICSM'14).

5.2. Cyclomatic complexity $\neq$ Lines of Code

It has long been believed that cyclomatic complexity of source code correlates linearly with lines of code (SLOC). After extensive study of a large corpus of Java source code, Davy Landman and Jurgen Vinju refuted this belief [34]. This provides a new landmark in how to assess and measure the quality of software. In short: cyclomatic complexity measures something different than lines of code.

5.3. Language-Parametric, Capture-Avoiding Program Transformation

Hygienic transformations are well-studied in the area of programming languages that feature (syntax) macros. For instance, in Scheme, macro expansion is guaranteed to not involuntarily capture existing bindings, or allow new bindings to be captured. Together with Sebastian Erdweg and Yi Dai, Tijs van der Storm designed a technique, "name-fix", that can be used to ensure hygiene in arbitrary program transformations, even when source and target language are completely different [24].

5.4. Memory Efficient Hash Tries

The hash trie data structure is a common part in standard collection libraries of JVM programming languages such as Clojure and Scala. It enables fast immutable implementations of maps, sets, and vectors, but it requires considerably more memory than an equivalent array-based data structure. Michael Steindorfer designed a product family of hash tries to generate specialized Java source code [29]. A preliminary experiment on the implementation of sets and maps shows that this technique leads to a median decrease of 55% in memory footprint for maps and 78% for sets.

5.5. Reflection without Remorse

A series of list appends or monadic binds for many monads performs algorithmically worse when it is left-associated. Continuation-passing style (CPS) is well-known to cure this severe dependence of performance on the association pattern. The advantage of CPS dwindles or disappears if we have to examine or modify the intermediate result of a series of appends or binds, before continuing the series. Such examination is frequently needed, for example, to control search in non-determinism monads. Atze van der Ploeg (together with Oleg Kiselyov) developed an alternative approach that is just as general as CPS but more robust: it makes series of binds and other such operations efficient regardless of the association pattern [30]. This solution solves previously undocumented, severe performance problems in iteratees, LogicT transformers, free monads and extensible effects.

5.6. General Parser Combinators

Parser combinators are a well-known approach to parsing where grammars are represented using (higher-order) functions. Unfortunately, parser combinators are commonly implemented using recursive descent parsing as the underlying algorithm. As a result, most parser combinators frameworks do not support left-recursive rules, and may exhibit exponential runtime performance due to backtracking. Anastasia Izmaylova and Ali Afroozeh developed "general parser combinators" (GPC) which do not suffer from these problems: all context-free grammars are supported (even ambiguous ones) and performance is worst-case cubic. As result, GPC combines the expressiveness and performance guarantees of general parsing algorithms like GLL and GLR with the flexibility and extensibility of parser combinators.

ATHENA Project-Team

6. New Results

6.1. Highlights of the Year

Maureen Clerc was awarded the PIERRE FAURRE Prize by the French Academy of Sciences. This award recognizes her outstanding contributions to the modelling and interpretation of electrical signals in the brain. The ceremony took place at the Institut de France on October 14th, 2013.

Emmanuel Caruyer was awarded the AFRIF Best PhD thesis award 2013 for his work "Q-space diffusion MRI: Acquisition and Signal Processing" performed under the direction of Rachid Deriche. He received the award thesis AFRIF 2013 during RFIA Conference held from June 30 to July 4, 2014 in Rouen.

Rachid Deriche was awarded the title of Honorary Doctor (honoris causa) from the University of Sherbrooke, Canada. This award recognises his achievements and contributions to image processing, computer vision and computational brain imaging. The title was awarded at the academic conferment ceremony held on September 20th, 2014 at the University of Sherbrooke.

Théo Papadopoulo has been promoted to the position of Research Director Class 2, starting from October 1st, 2014.

6.2. Modeling in Diffusion MRI

6.2.1. Non-Negative Spherical Deconvolution (NNSD) for estimation of fiber Orientation Distribution Function in single-/multi-shell diffusion MRI

Participants: Jian Cheng [University of North Carolina at Chapel Hill,USA], Tianzi Jiang [LIAMA, China], Shen Dinggang [University of North Carolina at Chapel Hill,USA], Yap Pew-Thian [University of North Carolina at Chapel Hill,USA], Rachid Deriche.

Spherical Deconvolution (SD) is commonly used for estimating fiber Orientation Distribution Functions (fODFs) from diffusion-weighted signals. Existing SD methods can be classified into two categories: 1) Continuous Representation based SD (CR-SD), where typically Spherical Harmonic (SH) representation is used for convenient analytical solutions, and 2) Discrete Representation based SD (DR-SD), where the signal profile is represented by a discrete set of basis functions uniformly oriented on the unit sphere. A feasible fODF should be non-negative and should integrate to unity throughout the unit sphere SS2. However, to our knowledge, most existing SH-based SD methods enforce non-negativity only on discretized points and not the whole continuum of SS2. Maximum Entropy SD (MESD) and Cartesian Tensor Fiber Orientation Distributions (CT-FOD) are the only SD methods that ensure non-negativity throughout the unit sphere. They are however computational intensive and are susceptible to errors caused by numerical spherical integration. Existing SD methods are also known to overestimate the number of fiber directions, especially in regions with low anisotropy. DR-SD introduces additional error in peak detection owing to the angular discretization of the unit sphere. This work proposes a SD framework, called Non-Negative SD (NNSD), to overcome all the limitations above. NNSD is significantly less susceptible to the false-positive peaks, uses SH representation for efficient analytical spherical deconvolution, and allows accurate peak detection throughout the whole unit sphere. We further show that NNSD and most existing SD methods can be extended to work on multi-shell data by introducing a three-dimensional fiber response function. We evaluated NNSD in comparison with Constrained SD (CSD), a quadratic programming variant of CSD, MESD, and an L1-norm regularized non-negative least-squares DR-SD. Experiments on synthetic and real single-/multi-shell data indicate that NNSD improves estimation performance in terms of mean difference of angles, peak detection consistency, and anisotropy contrast between isotropic and anisotropic regions.

This work has been published in [11].

6.2.2. Quantitative comparison of reconstruction methods for intra-voxel fiber recovery from diffusion MRI

Participants: Alessandro Daducci [LTS5, Ecole Polytech. Fed. de Lausanne (EPFL)], Maxime Descoteaux [SCIL Lab., Sherbrooke University], Michael Paquette [SCIL Lab., Sherbrooke University], Sylvain Merlet, Emmanuel Caruyer, Rachid Deriche.

Validation is arguably the bottleneck in the diffusion magnetic resonance imaging (MRI) community. This work evaluates and compares 20 algorithms for recovering the local intra-voxel fiber structure from diffusion MRI data and is based on the results of the “HARDI reconstruction challenge” organized in the context of the “ISBI 2012” conference. Evaluated methods encompass a mixture of classical techniques well known in the literature such as diffusion tensor, Q-Ball and diffusion spectrum imaging, algorithms inspired by the recent theory of compressed sensing and also brand new approaches proposed for the first time at this contest. To quantitatively compare the methods under controlled conditions, two datasets with known ground-truth were synthetically generated and two main criteria were used to evaluate the quality of the reconstructions in every voxel: correct assessment of the number of fiber populations and angular accuracy in their orientation. This comparative study investigates the behavior of every algorithm with varying experimental conditions and highlights strengths and weaknesses of each approach. This information can be useful not only for enhancing current algorithms and develop the next generation of reconstruction methods, but also to assist physicians in the choice of the most adequate technique for their studies.

This work has been published in [12]

6.2.3. Comparison of sampling strategies and sparsifying transforms to improve compressed sensing diffusion spectrum imaging

Participants: Michael Paquette [SCIL Lab., Sherbrooke University], Sylvain Merlet, Guillaume Gilbert [SCIL Lab., Sherbrooke University], Maxime Descoteaux [SCIL Lab., Sherbrooke University], Rachid Deriche.

Diffusion Spectrum Imaging enables to reconstruct the ensemble average propagator (EAP) at the expense of having to acquire a large number of measurements. Compressive sensing offers an efficient way to decrease the required number of measurements. The purpose of this work is to perform a thorough experimental comparison of three sampling strategies and six sparsifying transforms to show their impact when applied to accelerate compressive sensing-diffusion spectrum imaging. In this work, we propose a novel sampling scheme that assures uniform angular and random radial q-space samples. We also compare and implement six discrete sparse representations of the EAP and thoroughly evaluate them on synthetic and real data using metrics from the full EAP, kurtosis, and orientation distribution function. The discrete wavelet transform with Cohen–Daubechies–Feauveau 9/7 wavelets and uniform angular sampling in combination with random radial sampling showed to be better than other tested techniques to accurately reconstruct the EAP and its features. In conclusion, it is important to jointly optimize the sampling scheme and the sparsifying transform to obtain accelerated compressive sensing-diffusion spectrum imaging. Experiments on synthetic and real human brain data show that one can robustly recover both radial and angular EAP features while undersampling the acquisition to 64 measurements (undersampling factor of 4).

This work has been published in [16]

6.2.4. DSI 101: Better ODFs for Free

Participants: Michael Paquette [SCIL Lab., Sherbrooke University], Sylvain Merlet, Maxime Descoteaux [SCIL Lab., Sherbrooke University], Rachid Deriche.

Diffusion Spectrum Imaging (DSI) is a well established method to recover the Ensemble Average Propagator (EAP). The orientation distribution function (ODF) is computed from this discretized EAP and used for tractography. However, there are several important implementation considerations that are tossed aside in the literature and the publicly available softwares. We investigate all the real steps necessary to go from the DSI signal to the ODF and provide applicable recommendations that greatly improve the accuracy of the local orientation detected. These recommendations come ”free-of-charge” as they are applicable to all existing DSI data and do not require a significant increase in computation time.

This work has been published in [26]

6.2.5. Comparison between discrete and continuous propagator indices from Cartesian q-space DSI sampling

Participants: Mauro Zucchelli [Dpt of Computer Science, University of Verona], Eleftherios Garyfallidis [SCIL Lab., Sherbrooke University], Michael Paquette [SCIL Lab., Sherbrooke University], Maxime Descoteaux [SCIL Lab., Sherbrooke University], Gloria Menegaz [Dpt of Computer Science, University of Verona], Sylvain Merlet.

DSI is often considered the state-of-the-art technique to analyze q-space measurements sampled from a Cartesian grid. The 3D fast Fourier transform is used to directly obtain a discrete version of the EAP (Ensemble Average Propagator). DSI was one of the first techniques used to infer complex fiber configurations as it allows resolving crossings. In principle, DSI also captures some radial information which,in theory, can be used to extract diffusion features of the EAP. However, a discrete propagator representation suffers from a limited frequency band, which makes infinite integration impossible. Hence, EAP derived indices 2,3 are problematic and quantitatively questionable, as one needs to artificially normalize and approximate the infinite integrals. Combined with the recent popularity of DSI in the Human Connectome Project, it is important to investigate the different angular and EAP indices that can be computed from these DSI datasets. In this work, we investigate alternatives to the discrete model-free approach of DSI and investigate the Simple Harmonic Oscillator based Reconstruction and Estimation 3 (SHORE) models based on the evaluation of (i) the orientation distri bution function (ODF) ; (ii) the return to the origin probability (RTOP) and (iii) the mean square displacement (MSD).

This work has been published in [33]

6.2.6. Odf Maxima Computation Using Hill Climbing Algorithm

Participants: Makhlouf Laouchedi [USTHB, Algeria], Thinhinane Megherbi [USTHB, Algeria], Linda Oulebsir-Boumghar [USTHB, Algeria], Rachid Deriche.

Methods like Diffusion Spectrum Imaging (DSI), High Angular Resolution Diffusion Imaging (HARDI) and the High Order Tensor techniques have been proposed to reconstruct specific functions like the Orientation Distribution Function (ODF) whose maxima correspond to the directions of the multiple fibers. In this work, we are interested to extract all the crossing fibers characterized as the maxima of the Orientation Distribution Function (ODF). A Hill Climbing algorithm based approach has been developed and implemented to efficiently and accurately extract all the fibers. Promising experimental results obtained with synthetic and real data illustrate the potential of the technique.

This work has been published in [24]

6.2.7. Greedy NNLS: Fiber Orientation Distribution from Non-Negatively Constrained Sparse Recovery

Participants: Aurobrata Ghosh, Rachid Deriche.

In this work, we validated experimentally the merits of the Non-Negative Least Squares (NNLS) for the constrained sparse recovery of the Fiber Orientation Distribution (FOD) and compared it with classical l1-minimization. The FOD is a robust model for mapping crossing white matter fibers. However, its angular resolution depends on the spherical harmonic basis order, which can imply a large number of acquisitions. Further, it is necessary to compute the maxima of the FOD to derive the fiber directions. It is possible to kill the two proverbial birds with a single stone by using a non-negatively constrained sparse recovery FOD estimation with NNLS.

From our experiments, we confirmed results from literature to show that NNLS converges to highly sparse solutions which are correctly constrained, while l1-minimization is less sparse, contains negative solutions and is unstable with noisy data. Finally, we discussed the NLS algorithm and attributed the sparsity to its design, which mirrors the design of Orthogonal Matching Pursuit (OMP)

This work has been published in [22]

6.2.8. Crossing Fibers Detection with an Analytical High Order Tensor Decomposition

Participants: Thinhinane Megherbi [USTHB, Algeria], Mouloud Kachouane [USTHB, Algeria], Linda Oulebsir-Boumghar [USTHB, Algeria], Rachid Deriche.

Diffusion magnetic resonance imaging (dMRI) is the only technique to probe in vivo and noninvasively the fiber structure of human brain white matter. Detecting the crossing of neuronal fibers remains an exciting challenge with an important impact in tractography. In this, we tackle this challenging problem and propose an original and efficient technique to extract all crossing fibers from diffusion signals. To this end, we start by estimating, from the dMRI signal, the so-called Cartesian tensor fiber orientation distribution (CT-FOD) function, whose maxima correspond exactly to the orientations of the fibers. The fourth order symmetric positive definite tensor that represents the CT-FOD is then analytically decomposed via the application of a new theoretical approach and this decomposition is used to accurately extract all the fibers orientations. Our proposed high order tensor decomposition based approach is minimal and allows recovering the whole crossing fibers without any a priori information on the total number of fibers. Various experiments performed on noisy synthetic data, on phantom diffusion, data and on human brain data validate our approach and clearly demonstrate that it is efficient, robust to noise and performs favorably in terms of angular resolution and accuracy when compared to some classical and state-of-the-art approaches.

This work has been published in [15] and [34].

6.2.9. Complete set of Invariants of a 4th order tensor: the 12 tasks of HARDI from Ternary Quartics

Participants: Théodore Papadopoulo, Auro Ghosh, Rachid Deriche.

In this work, we presented a simple and systematic method to compute a functionally complete set of invariants of a non-negative 3D 4th order tensor with respect to 3D rotations. Intuitively, this transforms the tensor's non-unique ternary quartic (TQ) decomposition (from Hilbert's theorem) to a unique canonical representation independent of orientation.

Invariants play a crucial role in diffusion MRI. In DTI (2nd order tensors), invariant scalars (FA, MD...)have been successfully used in clinical applications. But DTI has limiations and HARDI models (e.g. 4th order tensors) have been proposed instead. These, however, lack invariant features and computing them systematically is challenging.

The invariants we propose, can be computed from two simple reduction steps, which first reduce an orthogonal class and then a rotation transform class of equivalent representations from the TQ coefficients. The resulting invariants are, by construction, (1) functionally complete, (2) functionally irreducible (if desired), (3) computationally efficient and (4) reversible – or mappable to the TQ coefficients or shape. These were the novelties of our contribution in comparison to prior work.

This work has been published in [25]

6.2.10. Fiber Orientation Distribution from Non-Negative Sparse Recovery

Participants: Thinhinane Megherbi [USTHB, Algeria], Auro Ghosh, Linda Oulebsir-Boumghar [USTHB, Algeria], Rachid Deriche.

In this work, we tested our non-negatively constrained sparse recovery algorithm for estimating the FOD on single shell phantom data provided by the ISBI'2014 challenge. We used the NNLS algorithm to estimate high order FODs (24th order) from just 20, 30 and 60 gradient directions and for various b-values of 1000, 2000, and 3000.

From the results, which are yet to be published, but can be viewed in their preliminary form online, it is clear that amongst the single shell algorithms, ours was good at fitting the signal and estimating the number of compartments. It performed well even with as low as 20 gradient acquisitions. Its major shortcoming was in underestimating the crossing angle and this needs to be improved upon.

This work has been published in [35]

6.2.11. How to get more out of a clinically feasable 64 gradient dMRI acquisition: Multi-Shell versus Single-Shell

Participants: Rutger H.j Fick, Mario Zuccheli [Dpt of Computer Science, University of Verona], Gabriel Girard [SCIL Lab., Sherbrooke University], Maxime Descoteaux [SCIL Lab., Sherbrooke University], Gloria Menegaz [Dpt of Computer Science, University of Verona], Rachid Deriche.

For clinical applications the number of diffusion MRI (dMRI) samples that can be obtained is often limited by scanner time and patient comfort. For this reason one often uses short scanning protocols that acquire just 32 or 64 gradient directions using a single b-value to obtain diffusion measures such as the fractional anisotropy from Diffusion Tensor Imaging (DTI) or to estimate the white matter orientation using Constrained Spherical Deconvolution (CSD). Using 3D-SHORE and MAP-MRI, we show that by spreading the same number of dMRI samples over different b-shells (sampling angularly and radially) we can estimate not only the directionality of the white matter using the ODF, but also the radially dependent higher order diffusion measures that SHORE and MAP-MRI provide. This approach lends itself well for situations where acquisition time is limited, and is therefore particularly well suited for clinical applications.

This work has been submitted to ISMRM'2015.

6.3. From dMRI to Fiber Pathways

6.3.1. Towards quantitative connectivity analysis: reducing tractography biases

Participants: Gabriel Girard [SCIL Lab., Sherbrooke University], Kevin Whittingstall [SCIL Lab., Sherbrooke University], Maxime Descoteaux [SCIL Lab., Sherbrooke University], Rachid Deriche.

Diffusion MRI tractography is often used to estimate structural connections between brain areas and there is a fast-growing interest in quantifying these connections based on their position, shape, size and length. However, a portion of the connections reconstructed with tractography is biased by their position, shape, size and length. Thus, connections reconstructed are not equally distributed in all white matter bundles. Quantitative measures of connectivity based on the streamline distribution in the brain such as streamline count (density), average length and spatial extent (volume) are biased by erroneous streamlines produced by tractography algorithms. In this work, solutions are proposed to reduce biases in the streamline distribution. First, we propose to optimize tractography parameters in terms of connectivity. Then, we propose to relax the tractography stopping criterion with a novel probabilistic stopping criterion and a particle filtering method, both based on tissue partial volume estimation maps calculated from a T1-weighted image. We show that optimizing tractography parameters, stopping and seeding strategies can reduce the biases in position, shape, size and length of the streamline distribution. These tractography biases are quantitatively reported using in-vivo and synthetic data. This is a critical step towards producing tractography results for quantitative structural connectivity analysis.

This work has been published in [13]

6.3.2. Choosing tractography parameters to improve connectivity mapping

Participants: Gabriel Girard [SCIL Lab., Sherbrooke University], Kevin Whittingstall [SCIL Lab., Sherbrooke University], Maxime Descoteaux [SCIL Lab., Sherbrooke University], Rachid Deriche.

Diffusion-weighted imaging (DWI) is often used as a starting point for in vivo white matter (WM) connectivity to reconstruct potential WM pathways between brain areas. Tractography algorithms have many parameters which can influence reconstruction and connectivity. Various choices of parameters have been proposed. But how to choose the best set of parameters ? In this study, we varied three critical parameters while monitoring connectivity score using the Tractometer evaluation system on the International Symposium on Biomedical Imaging (ISBI) Challenge synthetic dataset. The three parameters were: The maximum deviation angle between two consecutive tractography steps (this addresses the hypothesis of smoothness of the WM pathways), the spherical function (SF) threshold (this aims at removing noisy propagation directions during the tractography process) and the initial SF threshold (this aims at removing initial noise at the seeds and to start tractography in a good tangent direction to the WM bundle).

This work has been published in [20]

6.3.3. Connectivity directionally-encoded color map: a streamline-based color mapping

Participants: Gabriel Girard [SCIL Lab., Sherbrooke University], Kevin Whittingstall [SCIL Lab., Sherbrooke University], Maxime Descoteaux [SCIL Lab., Sherbrooke University], Rachid Deriche.

In this work, we provide a novel method to map streamlines in a color image, which can be generated from any set of streamlines. We show that this novel orientation color-coded map based on streamline tractography can improve connectivity analysis.

This work has been published in [23]

6.4. From dMRI to Microstructures Recovery

6.4.1. NMR characterization of cylinder radii distribution using a SHORE-based regularization method.

Participants: Gonzalo Sanguinetti, Matt Hall [Centre for Medical Image Computing, Dept. Computer Science, UCL], Daniel Alexander [Centre for Medical Image Computing, Dept. Computer Science, UCL], Rachid Deriche.

In this work, we are interested in retrieving information about the axon diameter distributions in white matter fiber bundles using NMR, which are commonly modelled as ensembles of parallel cylinders. We add regularization to the 1D-SHORE basis which results in more stable characterization of diameter distributions. To validate the method, we simulate NMR signals using the open source toolkit CAMINO. The results illustrate the enhanced estimation accuracy given by the regularization and provide an alternative validation of the SHORE based method.

This work has been published in [30]

6.4.2. Mapping Average axon diameters under long diffusion time

Participants: Gonzalo Sanguinetti, Rachid Deriche.

This work proposes an original method to recover axon diameter distribution (ADD) parameters using nuclear magnetic resonance. White matter (WM) is modelled as a bi-compartmental medium composed of an intra axonal space where the diffusion is restricted and an extra axonal space where diffusion is hindered. Under the assumption of long diffusion time, we provide a novel and efficient model for the component of the signal due to the restricted part. This technique might be interpreted as an interesting simplification of the AxCaliber framework, which leads to a simpler model and an extremely faster acquisition protocol. To test and validate our method, we use the open-source toolkit Camino for computing Monte-Carlo simulations of NMR data and model the WM as 3D cubic environments, formed by parallel cylinders with gamma distributed radii. Promising experimental results illustrate the potential of the proposed method.

This work has been published in [29] and [28]

6.4.3. Magnitude and complex based diffusion signal reconstruction

Participants: Marco Pizzolato, Timothe Boutelier [Olea Medical, La Ciotat], Rachid Deriche.

In Diffusion Weighted Magnetic Resonance Imaging (DW-MRI) the modeling of the magnitude signal is complicated by the Rician distribution of the noise. It is well known that when dealing instead with the complex valued signal, the real and imaginary parts are affected by Gaussian distributed noise and their modeling can thus benefit from any estimation technique suitable for this noise distribution. We present a quantitative analysis of the difference between the modeling of the magnitude diffusion signal and the modeling in the complex domain. The noisy complex and magnitude diffusion signals are obtained for a physically realistic scenario in a region close to a restricting boundary. These signals are then fitted with the Simple Harmonic Oscillator based Reconstruction and Estimation (SHORE) bases and the reconstruction performances are quantitatively compared. The noisy magnitude signal is also fitted by taking into account the Rician distribution of the noise via the integration of a Maximum Likelihood Estimator (MLE) in the SHORE. We discuss the performance of the reconstructions as function of the Signal to Noise Ratio (SNR) and the sampling resolution of the diffusion signal. We show that fitting in the complex domain generally allows for quantitatively better signal reconstruction, also with a poor SNR, provided that the sampling resolution of the signal is adequate. This applies also when the reconstruction is compared to the one performed on the magnitude via the MLE.

This work has been published in [27]

6.4.4. Extracting a biomarker for the mean cross-sectional area from the ODF

Participants: Rutger H.j Fick, Gonzalo Sanguinetti, Rachid Deriche.

Finding new biomarkers related to the microstructure of white matter (WM) is an active area of research in the MRI community. As opposed to the usual MRI markers such as fractional anisotropy (FA), these biomarkers provide a closer insight on the tissue structure. We introduce a new microstructure based biomarker that is related to the axon diameter distribution (ADD) and can be obtained with a q-space imaging technique like DSI or MAP . This feature is related with the nature and purpose of WM paths in bothnormal and pathological conditions and is obtained from the Orientation Distribution Function (ODF) as twice its maximum value. We show that this value is related with the mean cross-sectional area (MCSA) of an ensemble of parallel axons. The same geometric feature was proposed as a scalar index of microstructure, but was not related to the ODF. In this work we give the formal relation between this microstructure feature and the ODF, and validate it using state-of-the-art numerical simulations.

This work has been published in [18].

6.4.5. An Analytical 3D Laplacian Regularized SHORE Basis and Its Impact on EAP reconstruction and Microstructure Recovery

Participants: Rutger H.j Fick, Demian Wassermann, Gonzalo Sanguinetti, Rachid Deriche.

In diffusion MRI, the reconstructed Ensemble Average Propagator (EAP) from the diffusion signal provides detailed insights on the diffusion process and the underlying tissue microstructure. Recently, the Simple Harmonic Oscillator based Reconstruction and Estimation (SHORE) basis was proposed as a promising method to reconstruct the EAP. However, the fitting of the basis is sensitive to noise. To solve this we propose to use the Laplacian of the SHORE basis as a natural regularization functional. We provide the derivation of the Laplacian functional and compare its effect on EAP reconstruction with that of separated regularization of the radial and angular parts of the SHORE basis. To find optimal regularization weighting we use generalized cross-validation and validate our method quantitatively on synthetic and qualitatively on human data from the Human Connectome Project. We show that Laplacian regularization provides more accurate estimation of the signal and EAP based microstructural measures.

This work has been published in [19]

6.4.6. Using 3D-SHORE and MAP-MRI to obtain both Tractography and Microstructural Contrasts from a Clinical DMRI Acquisition

Participants: Rutger H.j Fick, Mario Zuccheli [Dpt of Computer Science, University of Verona], Gabriel Girard [SCIL Lab., Sherbrooke University], Maxime Descoteaux [SCIL Lab., Sherbrooke University], Gloria Menegaz [Dpt of Computer Science, University of Verona], Rachid Deriche.

Diffusion MRI (dMRI) is used to characterize the directional- ity and microstructural properties of brain white matter (WM) by measuring the diffusivity of water molecules. In clinical practice the number of dMRI samples that can be obtained is limited, and one often uses short scanning protocols that ac- quire just 32 to 64 different gradient directions using a single gradient strength (b-value). Such 'single shell' scanning protocols restrict one to use methods that have assumptions on the radial decay of the dMRI signal over different b-values, which introduces estimation biases. In this work, we show, that by simply spreading the same number of samples over multiple b-values (i.e. multi-shell) we can accurately estimate both the WM directionality using 3D-SHORE and characterize the radially dependent diffusion microstructure measures using MAP-MRI. We validate our approach by undersampling both noisy synthetic and human brain data of the Human Connectome Project, proving this approach is well-suited for clinical applications.

This work has been submitted to ISBI'2015.

6.4.7. Laplacian-Regularized MAP-MRI Improving Axonal Caliber Estimation

Participants: Rutger H.j Fick, Demian Wassermann, Gonzalo Sanguinetti, Rachid Deriche.

In diffusion MRI, the accurate description of the entire diffusion signal from sparse measurements is essential to enable the recovery of microstructural information of the white matter. The recent Mean Apparent Propagator (MAP)-MRI basis is especially well suited for this task, but the basis fitting becomes unreliable in the presence of noise. As a solution we propose a fast and robust analytic Laplacian regularization for MAP-MRI. Using both synthetic diffusion data and human data from the Human Connectome Project we show that (1) MAP-MRI has more accurate microstructure recovery com- pared to classical techniques, (2) regularized MAP-MRI has lower signal fitting errors compared to the unregularized approach and a positivity constraint on the EAP and (3) that our regularization improves axon radius recovery on human data.

This work has been submitted to ISBI'2015.

6.4.8. A Unifying Framework for Spatial and Temporal Diffusion in Diffusion MRI

Participants: Rutger H.j Fick, Demian Wassermann, Marco Pizzolato, Rachid Deriche.

We propose a novel framework to simultaneously represent the diffusion-weighted MRI (dMRI) signal over diffusion times, gradient strengths and gradient directions. Current frameworks such as the 3D Simple Harmonic Oscillator Reconstruction and Estimation basis (3D- SHORE) only represent the signal over the spatial domain, leaving the temporal dependency as a fixed parameter. However, microstructure- focused techniques such as Axcaliber and ActiveAx provide evidence of the importance of sampling the dMRI space over diffusion time. Up to now there exists no generalized framework that simultaneously models the dependence of the dMRI signal in space and time. We use a functional basis to fit the 3D+t spatio-temporal dMRI signal, similarly to the 3D-SHORE basis in three dimensional 'q-space'. The lowest order term in this expansion contains an isotropic diffusion tensor that characterizes the Gaussian displacement distribution, multiplied by a negative exponential. We regularize the signal fitting by minimizing the norm of the analytic Laplacian of the basis. The continuous 3D+t signal representation can provide new insights on the anomalous nature of the dMRI signal in human tissues, i.e., when mean-squared molecular displacements varies slower than linearly with the diffusion time. From the fitting one can also estimate the axon radius distribution parameters along any direction using approaches similar to AxCaliber. We validate our technique on synthetic data generated using the theoretical model proposed by Callaghan et al. We show that our method is robust to noise and can accurately describe the restricted spatio-temporal signal decay originating from tissue models such as cylindrical pores. Moreover, we apply our method on real data from an ActiveAx acquisition. Overall our approach allows to represent the complete 3D+t dMRI signal which should prove helpful in understanding normal and pathologic nervous tissue.

This work has been submitted to IPMI'2015.

6.4.9. *Fast and Robust EAP reconstruction using a Laplacian Regularized SHORE basis and its Impact on Microstructure Recovery*

Participants: Rutger H.j Fick, Demian Wassermann, Emmanuel Caruyer [SBIA, University of Pennsylvania Medical School], Rachid Deriche.

In diffusion MRI, the reconstructed Ensemble Average Propagator (EAP) from the diffusion signal provides detailed insights on the diffusion process and the underlying tissue microstructure. Recently, the 3D Simple Harmonic Oscillator based Reconstruction and Estimation (3D-SHORE) basis was proposed as a promising method to reconstruct the EAP. However, the fitting of the basis is sensitive to noise. To solve this we propose to use the Laplacian of the SHORE basis as a natural regularization functional. We provide the derivation of the Laplacian functional and compare its effect on EAP reconstruction with that of separated regularization of the radial and angular parts of the SHORE basis and imposing positive-definiteness in the estimation of the EAP. We validate our method on phantom data with known ground truth and on human data from the Human Connectome Project. We show that Laplacian regularization of the 3D-SHORE basis provides faster and more accurate estimation of the signal and EAP.

This work has been submitted to NeuroImage.

6.5. Functional and structural models analysis

6.5.1. *Analyzing Brain Plasticity in Math Learning Using Automated Dissection and Analysis of White Matter Tracts Through dMRI*

Participants: Dietsje Jolles [Stanford Medical School], Demian Wassermann, Ritika Chokhani [Stanford Medical School], Jennifer Richardson [Stanford Medical School], Caitlin Tenison [Stanford Medical School], Roland Bammer [Stanford Medical School], Lynn Fuchs [Vanderbit University], Kaustubh Supekar [Stanford Medical School], Vinod Menon [Stanford Medical School].

In a collaboration with Stanford Medical School, we explored longitudinal changes in white matter connectivity triggered by intensive math learning. Plasticity of white matter tracts is thought to be essential for cognitive development and academic skill acquisition in children. However, a dearth of high-quality diffusion tensor imaging (DTI) data measuring longitudinal changes with learning, as well as methodological difficulties in multi-time point tract identification have limited our ability to investigate plasticity of specific white matter tracts. With this contribution, we examined learning-related changes of white matter tracts innervating inferior parietal, prefrontal and temporal regions following an intense two-month math tutoring program. DTI data were acquired from 18 third grade children, both before and after tutoring. A novel fiber tracking algorithm based on a White Matter Query Language (WMQL) was used to identify three sections of the superior longitudinal fasciculus (SLF) linking frontal and parietal (SLF-FP), parietal and temporal (SLF-PT) and frontal and temporal (SLF-FT) cortices, from which we created child-specific probabilistic maps. The SLF-FP, SLF-FT, and SLF-PT tracts identified with the WMQL method were highly reliable across the two time points and showed close correspondence to tracts previously described in adults. Notably, individual differences in behavioral gains after two months of tutoring were specifically correlated with plasticity in the left SLF-FT tract. Our results extend previous findings of individual differences in white matter integrity, and provide important new insights into white matter plasticity related to math learning in childhood. More generally, our quantitative approach will be useful for future studies examining longitudinal changes in white matter integrity associated with cognitive skill development.

This work has been published in [14].

6.5.2. *Quantifying Uncertainty in Diffeomorphic Medical Landmark Registration*

Participants: Demian Wassermann, Matt Toew [Harvard Medical School - Brigham and Women's Hospital], Marc Niethammer [University of North Carolina at Chapel Hill], William Wells Iii [Harvard Medical School - Brigham and Women's Hospital, MIT].

In a collaboration with Harvard Medical School, the Brigham and Women's Hospital, MIT and the University of North Carolina at Chapel Hill, we proposed a novel mathematical framework to represent uncertainty in diffeomorphic registration techniques. Particularly, we introduced a novel mathematical framework for representing uncertainty in large deformation diffeomorphic image registration. The Bayesian posterior distribution over the deformations aligning a moving and a fixed image is approximated via a variational formulation. A stochastic differential equation (SDE) modeling the deformations as the evolution of a time-varying velocity field leads to a prior density over deformations in the form of a Gaussian process. This permits estimating the full posterior distribution in order to represent uncertainty, in contrast to methods in which the posterior is approximated via Monte Carlo sampling or maximized in maximum a-posteriori (MAP) estimation. The framework was demonstrated in the case of landmark-based image registration, including simulated data and annotated pre and intra-operative 3D images. This type of registration can be extended to several anatomical objects such as white matter tracts represented as streamlines.

This work has been published in [32].

6.5.3. Group Comparisons on White Matter Tracts in Native Space

Participants: Eleftherios Garyfallidis [University of Sherbrooke], Demian Wassermann, Maxime Descoteaux [University of Sherbrooke].

Let us suppose that we want to study specific fiber bundles in different subjects. The common approach would be to use a voxel-wise analyses which will warp scalar volumes in a common space, e.g. MNI space, and show how every subject differentiates from an average template. However, we know that with averaging and warping much of the specific information about the individual subjects' differences is lost. In this work, we provide a solution to this problem by using local streamline registration of specific bundles from different subjects. We show that with this new method we can keep track of the differences from every subject to every other subject in our group study.

This study was performed in collaboration with the SCIL lab of Sherbrook University within the framework of the Brain Connectivities Associate Team and published in [21].

6.5.4. Perfusion Deconvolution via SHORE and Laplacian Regularization

Participants: Marco Pizzolato, Auro Ghosh, Timothé Boutelier [Olea Medical, La Ciotat], Rachid Deriche.

Perfusion imaging comprehensively refers to the recovery of parameters of interest which are related to the passage of blood in the parenchyma (i.e. the functional part) of a tissue. The amount of perfusion is related to both the functionality of the parenchyma and its level of activity. By means of imaging techniques such as Dynamic Susceptibility Contrast MRI it is possible, in each voxel, to measure the tissue concentration $Ct(t)$ of a tracer injected before the scanning in the vascular system. According to the indicator dilution theory1 this is related to the concentration measured in an arterial region $Ca(t)$ described by a convolution with $R(t)$ that is the unknown residue function expressing the remaining time-dependent tracer quantity in the voxel. Historically $R(t)$ is obtained exploiting the convolution theorem $R(t) = FT - 1FT[Ct(t)]/FT[Ca(t)]$. However deconvolution is an ill-posed problem making this method very sensitive to noise. Many regularization techniques have been proposed but among all the most adopted technique is truncated Singular Value Decomposition, tSVD. However tSVD is known to underestimate an important perfusion parameter that is the blood flow BF, which can be computed as the maximum peak of the recovered $R(t)$. In this work we propose to use the Simple Harmonic Reconstruction and Estimation framework (SHORE) to estimate $R(t)$ in order to obtain a better parameter estimation. We regularize SHORE using Laplacian regularization. We compare the results with tSVD.

This work has been submitted to ISMRM 2015.

6.5.5. Perfusion MRI Deconvolution with Delay Estimation and Non-negativity Constraints

Participants: Marco Pizzolato, Auro Ghosh, Timothé Boutelier [Olea Medical, La Ciotat], Rachid Deriche.

Perfusion MRI deconvolution aims to recover the time-dependent residual amount of indicator (residue function) from the measured arterial and tissue concentration time-curves. The deconvolution is complicated by the presence of a time lag between the measured concentrations. Moreover the residue function must be non-negative and its shape may become non-monotonic due to dispersion phenomena. We introduce Modified Exponential Bases (MEB) to perform deconvolution. The MEB generalizes the previously proposed exponential approximation (EA) by taking into account the time lag and introducing non-negativity constraints for the recovered residue function also in the case of non-monotonic dispersed shapes, thus overcoming the limitation due to the non-increasing assumtion of the EA. The deconvolution problem is solved linearly. Quantitative comparisons with the widespread block-circulant Singular Value Decomposition show favorable results in recovering the residue function.

This work has been submitted to ISBI 2015.

6.6. Forward and Inverse Problems in MEEG

6.6.1. FindSource3D - Source Localization Using Rational Approximation on Plane Sections

Participants: Todor Jordanov [BESA GmbH, Germany], Jean-Paul Marmorat [École des Mines ParisTech, Sophia Antipolis], Maureen Clerc, Juliette Leblond, Andre Waelkens [BESA GmbH, Germany], Théodore Papadopoulo.

A new method for EEG source localization based on rational approximation techniques in the complex plane was suggested. The method is used in the context of a nested sphere head model, in combination with a cortical mapping procedure [51]. This method was shown to perform perfectly for numerical simulations without noise but its performance with respect to different signal-to-noise ratios (SNRs), to different number of sources and to real EEG data was not investigated until now. The method, formally called FindSource3D (FS3D), is evaluated with data simulations and a real EEG data set.

This work has been published in [40].

6.6.2. Diffusion Magnetic Resonance information as a regularization term for MEG/EEG inverse problem

Participants: Brahim Belaoucha, Anne-Charlotte Philippe, Maureen Clerc, Théodore Papadopoulo.

Several regularization terms are used to constrain the Magnetoencephalography (MEG) and the Electroencephalography (EEG) inverse problem. It has been shown that the brain can be divided into several regions with functional homogeneity inside each one of them. To locate these regions, we use the structural information coming from the diffusion Magnetic Resonance (dMRI) and more specifically, the anatomical connectivity of the distributed sources computed from dMRI. To investigate the importance of the dMRI in the source reconstruction, this work compares the solutions based on dMRI-based parcellation to random parcellation.

This work has been published in [37].

6.6.3. Dictionary learning for multitrial datasets

Participants: Maureen Clerc, Sebastian Hitziger, Théodore Papadopoulo.

Following the path opened with the Consensus matching Pursuit method (CMP) [48], we continue our endeavour to avoid signal averaging using directly the raw signal with the assumption that events of interest are those that repeat in each trial. Towards such a goal, and to improve the simple dictionary used in CMP, we have adapted dictionary learning methods to multitrial bio-electric signals, by explicitly implementing jitter invariance [62]. This allows for a much more detailed data-driven description of events. For example, using local field potential signals of chemically induced spikes (in a rat model), we have been able to distinguish several spike shapes which show some coherence in time. The method has been recently extended to detect spike events in continuous signals (i.e. not organized in epochs). While it requires a good signal to noise ratio, the method is very general and has also been used for various other signal types (see section 6.7).

This work has been published in [39].

6.7. Coupling functional and structural models

6.7.1. Propagation of epileptic spikes revealed by diffusion-based constrained MEG source reconstruction

Participants: Anne-Charlotte Philippe, Théodore Papadopoulo, Christian Bénar [Hospital "La Timone", Marseille], Jean-Michel Badier [Hospital "La Timone", Marseille], Maureen Clerc, Rachid Deriche.

In this work, we study the propagation of an epileptic spike (from single event data). As in the two previous sections, a cortex parcellation is performed using structural information coming from diffusion MRI Then, a MEG inverse problem is defined on a parcellated source space which imposes constant activity on each parcel. This inverse problem is applied separately for measurements obtained in a given time range. The most active parcels over the time range are located and their time course are displayed. This allowed the study of the propagation of an epileptic spike via those active parcels. Results on real data shows varying spatial propagations of an epileptic spike for the same subject.

This work has been published in [41].

6.7.2. Using diffusion MRI information in the Maximum Entropy on Mean framework to solve MEG/EEG inverse problem

Participants: Brahim Belaoucha, Jean-Marc Lina [Centre de Recherches Mathématique, Montréal], Maureen Clerc, Anne-Charlotte Philippe, Christophe Grova [McGill University], Théodore Papadopoulo.

Magnetoencephalography (MEG) and Electroencephalography (EEG) inverse problem is well-known to require regularization in order to avoid ill-posedness. Usually, regularization is based on mathematical criteria (minimum norm, ...). Physiologically, the brain is organized in functional parcels and imposing a certain homogeneity of the activity within these parcels was proven to be an efficient way to analyze the MEG/EEG data. The parcels information can be computed from diffusion Magnetic Resonances Imaging (dMRI) by grouping together source positions shared the same connectivity profile (computed as tractograms from diffusion images). In this work, three parcel-based inverse problem approaches have been tested. The first two approaches are based on minimum norm with added regularization terms to account for the parcel information. They differ by the use of a hard/soft constraint in the way they impose that the activity is constant within each parcel [74]. The third approach is based on the Maximum Entropy on Mean (MEM) framework [42]. It models source activity with a random variable and parcels are also used as a regularization. Several tests have been conducted with synthetic and real data that encompass the MEG/EEG and the diffusion magnetic resonance signals to compare these three approaches in terms of active region-detection accuracy.

This work has been published in [36].

6.8. Brain Computer Interfaces

6.8.1. CoAdapt P300 speller: optimized flashing sequences and online learning

Participants: Maureen Clerc, Théodore Papadopoulo, Loïc Mahé.

Our work in Brain Computer Interfaces was centered around the visual P300 speller system: a virtual keyboard allowing to type words by detecting the P300 wave, an automatic deflection of the central component of the electric potential, occurring approximately 300 ms after the apparition of an intermittent and rare event, on which the user's attention is focussing. The idea behind the P300 speller is very simple: the system displays series of stimuli (flashes), over the keyboard elements, and detects whether or not the EEG recorded after each flash contains a P300. Its implementation is not so simple, because of the low amplitude of the P300 compared to the background EEG, and of the inter-subject variability of this signal.

The advantage of this system is not to require any training on the part of the user. However the BCI system has to be trained to detect the P300 component from the background EEG: this is done through a calibration phase.

We developped a new method to reduce the calibration phase, with a transfer learning method called “mixture of experts” (MOE). The MOE classifier makes its predictions by averaging the decisions from a pre-recorded database of classifiers coming from other recording sessions (with other subjects) [31]. The decisions were made by using an evidence accumulation scheme, which updated the prediction at every flash of the keyboard [17].

Part of this work has been implemented in the software: CoAdapt P300 Stimulator.

6.8.2. P300 speller: clinical feasibility study with Amyotrophic Lateral Sclerosis

Participants: Maureen Clerc, Théodore Papadopoulo, Loïc Mahé, Asya Metelkina, Violaine Guy [Nice University Hospital], Claude Desnuelle [Nice University Hospital].

From September 2013 to July 2014, we were very involved in running an experiment with the Centre de Référence Maladies Neuromusculaires et SLA (CRMN/SLA) of Nice University Hospital. This study, partly funded by “Assocation pour la Recherche sur la Sclérose Latérale Amyotrophique”, was conducted on 20 patients, who routinely come to be examined at the hospital. Each patient came for 3 sessions where he/she was allowed to use the P300 speller, after being equipped with an electro-encephalography cap, and watching a video explaining the modus operandi of the P300 and on their role in the study.

The P300 speller system has been organized in a way to make it relatively easy to deploy in a clinical setting: it involves only one laptop, and requires limited intervention from the caregiver. The most intricate operation is to position the EEG headset and ensure a correct impedance (below 5 kΩ) for all electrodes.

Each session consisted of three blocks, after the initial calibration phase: a copy spelling task of two ten-letter words, a free spelling task of approximately twenty characters, and an optional block of free use of the system for writing. Finally, the patient was asked to answer a questionnaire. This study intends to investigate the feasibility of setting up and using the P300 speller, from an operational point of view at the hospital. Translational studies of this type are extremely important for the adaptation the BCI systems to the target patient populations, and a large-scale usability study for the P300 speller has never been done before in France.

6.8.3. BCI Challenge: A spell on you if you cannot detect errors!

Participants: Maureen Clerc, Théodore Papadopoulo, Jérémie Mattout [Centre de Recherche en Neurosciences de Lyon, INSERM], Emmanuel Maby [Centre de Recherche en Neurosciences de Lyon, INSERM].

We have proposed an international BCI Challenge on decoding Error Potential signals. The winners will be announced at the 7th International IEEE/EMBS Conference on Neural Engineering, Montpellier, in April 2015. The Challenge was open on the Kaggle platform on Nov 14, 2014 and will close on Feb 24, 2015, see: website. At this date the competition is still open, and it has so far attracted 212 participants forming 181 teams.

In the P300 speller paradigm (see above) and in other BCI where a discrete feedback can be presented to the user, the EEG evoked response to the feedback can be recorded and processed online in order to evaluate whether the item selection was correct or not. This decision, if reliable, could then be used to improve the BCI performance by implementing some error correction strategy. In this competition, participants are asked to submit an Error Potential detection algorithm, capable of detecting the erroneous feedbacks online and to generalize across subjects (transfer learning).

The data used in this competition was acquired in the scope of the CoAdapt ANR project.

ATLANMOD Project-Team

6. New Results

6.1. Model Quality

Our work aims to enhance the quality of the modeling activity in the context of software engineering and language engineering. This year, this has translated in the following results:

- A systematic review [16] of all formal verification approaches targeting the quality evaluation of software models to be used as the basis for future research on the topic and as a kind of reference comparison to compare new tools with existing ones.
- A complete description of our CSP-based approach for the verification of UML/OCL models (where both the uml constructs and OCL expressions are translated into a constraint satisfaction problem) [12]
- A new test data generation approach for Model Transformations that combines partitions and constraint analysis to try to mazimize the coverage of the generates tests [29]

6.2. Model Driven approach to mobile applications development

Cross-platform and multi-device design, implementation and deployment is a barrier for today's IT solution providers, especially SME providers, due to the high cost and technical complexity of targeting development to a wide spectrum of devices, which differ in format, interaction paradigm, and software architecture. Our work aims at exploiting the modern paradigm of Model-Driven Engineering and code generation to simplify multi-device development, reducing cost and development times, so as to increase the profit of SME solution providers and at the same time reduce the price and total cost of ownership for end-customers. In [22] we defined a Platform Independent Modeling language for mobile applications. The language has been defined as a mobile extension of an OMG standard called Interaction Flow Modeling Language (IFML). The research included also the development of an Eclipse-based modeling tool for mobile apps and the first prototypes of automatic code generators.

6.3. Security

Most companies information systems are composed by heterogeneous components responsible of hosting, creating or manipulating critical information for the day-to-day operation of the company. Securing this information is therefore one of their main concerns, more particularly specifying Access Control (AC) policies. However, the task of implementing an AC security policy (sometimes relying on several mechanisms) remains complex and error prone as it requires knowing low level and vendor-specific facilities. In this context, discovering and understanding which security policies are actually being enforced by the Information System (IS) becomes critical. Thus, the main challenge consists in bridging the gap between the vendor-dependent security features and a higher-level representation. This representation has to express the policies by abstracting from the specificities of the system components, allowing security experts to better understand the policy and to implement all related evolution, refactoring and manipulation operations in a reusable way.

In 2014, we have presented a Ph.D. thesis tackling the aforementioned problems. It proposes a model-driven automatic reverse engineering mechanism capable of analyzing deployed security aspects of components (e.g. concrete firewall configurations) to derive the abstract model (e.g. network security global policy) that is actually enforced. Once the model is obtained, it can be reconciled with the expected security directives, to check its compliance, can be queried to test consistency or used in a process of forward engineering to generate validated security configurations. This work also provides the first steps towards the integration of the diverse security policies extracted from the subsystems composing a complex Information System in a global security representation.

6.4. Model-Driven Document Engineering

As a result of a long-term collaboration of one of the AtlanMod team members with the ISSI research group at the Universitat Politècnica de València, we have participated in the publication of several works on the area of the Document Engineering. In this research line, we have applied the MDE methods and tools to the product-line-based generation of customized documents resulting in the so-called DPL methodology [0]. The Document Product Lines (DPL) approach, which we throughly describe in a journal publication [17], provides a framework for variable content document generation that follows an alternate path to the traditional variable document generation. DPL has been created with a twofold goal: first, to make creating variable content documents available to non-experts by including a domain engineering process previous to the document generation itself; and secondly, to enforce content reuse at domain level.

DPLFW is the main tool supporting the DPL methodology, and in the demonstrations track of the MODELS conference we showed all its capabilities. In addition to these contributions, we have published several works demonstrating the applicability of the DPL–DPLFW tandem in different domains, such as the development of executable emergency plans in crisis managent contexts [25], the development of learning objects in the e-learning field [32] and the generation of customized documents in e-Government solutions [33].

6.5. Reverse Engineering and Evolution

Model Driven Reverse Engineering (MDRE), and its applications such as software modernization, is a discipline in which model-driven development (MDD) techniques are used to treat legacy systems. During this year, Atlanmod has continued working actively on this research area. The main contributions are the following:

- In the context of the ARTIST FP7 project, the work has been continued on reusing (and extending accordingly) MoDisco and several of its components to provide the Reverse Engineering support required within the project. At conceptual-level, the MoDisco Model Discovery + Model Understanding overall two-step approach [11] has been published and promoted as an important part of the ARTIST migration methodology and process [18]. At tooling-level, several (MoDisco-based) model discovery components from Java and SQL have been developed and made available as part of the official ARTIST OS Release [0]. Directly related to some of these components, a promising work has been initiated on studying deeper the discovery of behavioral aspects of software and dealing with their further understanding based on the OMG FUML standard combined with different modeling techniques (transformation, slicing, etc.). Complementary work has also been performed in the context of the TEAP FUI project finishing by the end of this year. It concerns the related problem of data federation from heterogeneous sources in the domain of Enterprise Architecture. This has notably resulted in a prototype called EMF Views that can be practically used in such reverse engineering scenarios [36] and also in other cases to be further explored (cf. the MoNoGe FUI project dealing with (meta)model extension).
- In a web context, in a previous work we shown how to discover the schema which is implicit in JSON data. This year we built on that contribution to study how schemas coming from different JSON-based web APIs can be composed [24]. Thus, we presented an approach able to identify composition links between schemas of different APIs. This composition information plus the API schemas are used to render a graph where paths represent API compositions and are used to easily identify how to compose the APIs. For instance, we illustrated one application based on generating sequence diagrams from graph paths, where the diagram includes the API calls (and their corresponding parameters) that web developers have to perform in order to compose one or more APIs.
- In the context of our work around DSLs, we have been working on facilitating the definition of DSLs from existing APIs. Sometimes library developers prefer to provide their users with a DSL, instead of (or in addition to) an API. APIs and DSLs can be seen as alternative methods to access the library

[0] http://dpl.dsic.upv.es, only in Spanish
[0] http://www.artist-project.eu/tools-of-toolbox/193

functionalities, and are characterized by specific advantages. We therefore proposed a method to automatically analyze an existing object-oriented API and generate a DSL out of it. Our approach leverages on model-driven techniques to analyze and represent APIs at high-level of abstraction (i.e., as metamodels) which are later used to automatically generate the DSL components and the corresponding tooling, including parser, compiler and development environment. Developers can influence the DSL generation by editing the model-based API representation and by specifying design choices about the structure of the DSL to generate. A proof-of-concept implementation of the method has been developed, called *DSLit*, that is able to analyze Java APIs and generate textual DSLs.

- On the evolution side, we have been working on an approach to automatically resynchronize code-generation artefacts (in particular, model-to-text transformations) after changes on the target platform [28]

6.6. Scalability

The increasing number of companies embracing MDE methods and tools have exceeded the limits of the current model-based technologies, presenting scalability issues while facing the growing complexity of their data. Since further research and development is imperative in order to maintain MDE techniques as relevant as they are in less complex contexts, we have focused our research in three axes, (i) scalable persistence solutions, (ii) scalable model transformation engines, and (iii) testing of large scale distributed systems.

In [21], we lead the first open-set benchmark gathered from real-world cases to stress scalability issues in model transformation and query engines. This benchmark suite has been made public with a twofold goal: (i) to provide a reference benchmark suite to both the industry and the research community that can be used to compare and evaluate different technologies that may fulfill their needs; and (ii) to motivate the MDE community to be part of its development by allowing them to extend and contribute with additional cases not covered by the initial set.

On the other hand, we introduce Neo4EMF [20], a NoSQL database persistence framework based on Neo4j [0]. Neo4EMF provides light-weight on-demand loading and storage facilities for handling very large models. Additionally, we also show that Neo4EMF can handle the creation of very-large models without performing periodical saves manually.

In this paper [31], we argue that fUML may be leveraged to address the well-known interoperability issue between tools from different modeling platforms. This is done by providing a common execution language and by abstracting modeling frameworks into generic actions that perform elementary operations on models. User models can not only benefit from a unified execution semantics, but also modeling tools can benefit too. As a proof of concept, we show [37] how it can be applied to model transformation engines, in particular ATL. To this end, an prototype compiler from ATL to fUML has been built.

In [19], we present a model-based approach to define a dynamic oracle for checking global properties on distributed software. Our objective is to abstract relevant aspects of such systems into models by gathering data from different nodes and building a global view of the system, where properties are validated. These models are updated at runtime, by monitoring the corresponding distributed system. This process requires a distributed test architecture and tools for representing and validating global properties. To evaluate the ability of our approach, a real-scale experimental validation has been conducted.

[0] http://www.neo4j.org

AVALON Project-Team

6. New Results

6.1. Energy efficiency of large scale distributed systems

Participants: Laurent Lefèvre, Daniel Balouek Thomert, Eddy Caron, Radu Carpa, Ghislain Landry Tsafack Chetsa, Marcos Dias de Assunçao, Jean-Patrick Gelas, Olivier Glück, Jean-Christophe Mignot, François Rossigneux, Violaine Villebonnet.

6.1.1. Improving Energy Efficiency of Large Scale Systems without a priori Knowledge of Applications and Services

Unlike their hardware counterpart, software solutions to the energy reduction problem in large scale and distributed infrastructures hardly result in real deployments. At the one hand, this can be justified by the fact that they are application oriented. At the other hand, their failure can be attributed to their complex nature which often requires vast technical knowledge behind proposed solutions and/or thorough understanding of applications at hand. This restricts their use to a limited number of experts, because users usually lack adequate skills. In addition, although subsystems including the memory and the storage are becoming more and more power hungry, current software energy reduction techniques fail to take them into account. We propose a methodology for reducing the energy consumption of large scale and distributed infrastructures. Broken into three steps known as (i) phase identification, (ii) phase characterization, and (iii) phase identification and system reconfiguration; our methodology abstracts away from any individual applications as it focuses on the infrastructure, which it analyses the runtime behaviour and takes reconfiguration decisions accordingly.

The proposed methodology is implemented and evaluated in high performance computing (HPC) clusters of varied sizes through a Multi-Resource Energy Efficient Framework (MREEF). MREEF implements the proposed energy reduction methodology so as to leave users with the choice of implementing their own system reconfiguration decisions depending on their needs. Experimental results show that our methodology reduces the energy consumption of the overall infrastructure of up to 24% with less than 7% performance degradation. By taking into account all subsystems, our experiments demonstrate that the energy reduction problem in large scale and distributed infrastructures can benefit from more than "the traditional" processor frequency scaling. Experiments in clusters of varied sizes demonstrate that MREEF and therefore our methodology can easily be extended to a large number of energy aware clusters. The extension of MREEF to virtualized environments like cloud shows that the proposed methodology goes beyond HPC systems and can be used in many other computing environments.

6.1.2. Reservation based Usage for Energy Efficient Clouds: the Climate/Blazar Architecture

The FSN XLcloud project (cf Section 8.1) strives to establish the demonstration of a High Performance Cloud Computing (HPCC) platform based on OpenStask, that is designed to run a representative set of compute intensive workloads, including more specifically interactive games, interactive simulations and 3D graphics. XLcloud is based on OpenStack, and Avalon is contributing to the energy efficiency part of this project. We have proposed and brought our contribution to Climate, a new resource reservation framework for OpenStack, developed in collaboration with Bull, Mirantis and other OpenStack contributors. Climate allows the reservation of both physical and virtual resources, in order to provide a mono-tenancy environment suitable for HPC applications. Climate chooses the most efficient hosts (flop/W). This metric is computed from the CPU / GPU informations, mixed with real power consumption measurements provided by the Kwapi framework. The user requirements may be loose, allowing Climate to choose the best time slot to place the reservation. Climate has been improved with standby mode features, to shut down automatically the unused hosts. The first release of Climate was done in January 2014. Through the OpenStack process, Climate is now named Blazar.

6.1.3. Clustered Virtual Home Gateway (vHGW)

This result is a joint work between Avalon team (J.P. Gelas, L. Lefevre) and Addis Abeba University (M. Tsibie and T. Assefa). The customer premises equipment (CPE), which provides the interworking functions between the access network and the home network, consumes more than 80% of the total power in a wireline access network. In the GreenTouch initiative (cf Section 8.3), we aim at a drastic reduction of the power consumption by means of a passive or quasi-passive CPE. Such approach requires that typical home gateway functions, such as routing, security, and home network management, are moved to a virtual home gateway (vHGW) server in the network. In our first prototype virtual home gateways of the subscribers were put in LXC containers on a unique GNU/Linux server. The container approach is more scalable than separating subscribers by virtual machines. We demonstrated a sharing factor of 500 to 1000 virtual home gateways on one server, which consumes about 150 W, or 150 to 300 mW per subscriber. Comparing this power consumption with the power of about 2 W for the processor in a thick client home gateway, we achieved an efficiency gain of 5-10x. The prototype was integrated and demonstrated at TIA 2012 in Dallas. In our current work, we propose the Clustered vHGWs Data center architecture to yield optimal energy conservation through virtual machine's migration among physical nodes based on the current subscriber's service access state, while ensuring SLA respective subscribers. Thus, optimized energy utilization of the data center is assured without compromising the availability of service connectivity and QoS preferences of respective subscribers. The last prototype including those new features was integrated and demonstrated recently to the GreenTouch consortium members at Melbourne University.

6.1.4. Energy proportionality with heterogeneous computing resources

This work [16] focuses on improving energy proportionality of large scale virtualized environments. The main problem of such infrastructures is their high static costs due to high idle power consumption of idle servers. Our goal is to reach an infrastructure able to adapt its energy consumption to the current working load. Therefore we propose an original infrastructure composed of heterogeneous computing resources. We consider the heterogeneity at the level of the architecture, and we gather in our platform low power ARM processors together with powerful x86 servers. Around this infrastructure, we are developing a decisional framework to schedule applications on the architecture, or combination of architectures, most suitable to their current needs. The framework reacts dynamically to the resource needs evolutions by migrating the applications to the chosen destinations, and switching off unused nodes to save energy. We validate our scheduling policies by building a simulator based on a set of experimental inputs about power and performance hardware profiles and applications load profiles. This work is jointly done with IRIT Lab. (Toulouse) under the support of Inria Large Scale Initiative Hemera.

6.1.5. Energy efficient Core Networks

This work [11] seeks to improve the energy efficiency of backbone networks by providing an intra-domain Software Defined Network (SDN) approach to selectively turn off a subset of links. To do this, we change the status of router ports and transponders on the two extremities of a link. The status of these components is set to sleep mode whenever a link is not required to transfer data, and brought back to operational state when needed. We have analyzed the implementation issues of an energy-efficient SDN-based traffic engineering in core networks. We propose the STREETE framework (SegmenT Routing based Energy Efficient Traffic Engineering) that represents an online method to switch some links off/on dynamically according to the network load. We have implemented our proposed algorithms in the OMNET++ packet-based discrete event simulator. Experiments considering real network topologies (Germany50 and Ge´ant) and real dynamic traffic matrices allowed us to quantify the trade-off between energy saving and impact of our solution on network performance. As mean to reroute the traffic we use a promising new protocol, SPRING. This comes in contrast with other works, which use classical IP link weights changes or MPLS+RSVP-TE for this purpose. SPRING proved itself well suited for dynamic reconfiguration of the network. Experimental results show that the consumption of 44% of links can be reduced while preserving good quality of service.

6.1.6. Energy aware scheduling for multi data centers clouds

Our work tackles the challenge of improving the energy efficiency of server provisioning and workload management [17]. It introduces a metric allowing infrastructure administrators to specify their preferences between performance and energy savings. We describe a framework for resource management which provides control for informed and automated provisioning at the scheduler level while providing developers (administrator or end-user) with an abstract layer to implement aggregation and resource ranking based on contextual information such as infrastructure status, users' preferences and energy- related external events occurring over time. We integrate our solution in DIET which allows for managing heterogeneous nodes at the middleware layer. The evaluation is performed by means of simulations and real-life experiments on the GRID'5000 testbed. Results show improvements in energy efficiency with minimal impact on application and system performance. Implementation has been used within the industrial project Nu@ge in the context of a federation of modular datacenters.

6.2. Simulation of Large Scale Distributed Systems

Participants: Frédéric Desprez, Jonathan Rouzaud-Cornabas, Frédéric Suter.

6.2.1. Versatile, Scalable, and Accurate Simulation of Distributed Applications and Platforms

The study of parallel and distributed applications and platforms, whether in the cluster, grid, peer-to-peer, volunteer, or cloud computing domain, often mandates empirical evaluation of proposed algorithmic and system solutions via *simulation*. Unlike direct experimentation via an application deployment on a real-world testbed, simulation enables fully repeatable and configurable experiments for arbitrary hypothetical scenarios. Two key concerns are accuracy (so that simulation results are scientifically sound) and scalability (so that simulation experiments can be fast and memory-efficient). While the scalability of a simulator is easily measured, the accuracy of many state-of-the-art simulators is largely unknown because they have not been sufficiently validated. In this work we describe recent accuracy and scalability advances made in the context of the SIMGRID simulation framework. A design goal of SIMGRID is that it should be versatile, i.e., applicable across all aforementioned domains. We present quantitative results that show that SIMGRID compares favorably to state-of-the-art domain-specific simulators in terms of scalability, accuracy, or the trade-off between the two. An important implication is that, contrary to popular wisdom, striving for versatility in a simulator is not an impediment but instead is conducive to improving both accuracy and scalability.

6.2.2. Simulation of MPI Applications with Time-Independent Traces

Analyzing and understanding the performance behavior of parallel applications on parallel computing platforms is a long-standing concern in the High Performance Computing community. When the targeted platforms are not available, simulation is a reasonable approach to obtain objective performance indicators and explore various hypothetical scenarios. In the context of applications implemented with the Message Passing Interface, two simulation methods have been proposed, on-line simulation and off-line simulation, both with their own drawbacks and advantages. In this work we present an off-line simulation framework, i.e., one that simulates the execution of an application based on event traces obtained from an actual execution. The main novelty of this work, when compared to previously proposed off-line simulators, is that traces that drive the simulation can be acquired on large, distributed, heterogeneous, and non-dedicated platforms. As a result the scalability of trace acquisition is increased, which is achieved by enforcing that traces contain no time-related information. Moreover, our framework is based on an state-of-the-art scalable, fast, and validated simulation kernel.

6.2.3. Adding Storage Simulation Capacities to the SimGrid Toolkit

For each kind of distributed computing infrastructures, i.e., clusters, grids, clouds, data centers or supercomputers, storage is a essential component to cope with the tremendous increase in scientific data production and the ever-growing need for data analysis and preservation. Understanding the performance of a storage subsystem or dimensioning it properly is an important concern for which simulation can help by allowing for fast, fully repeatable, and configurable experiments for arbitrary hypothetical scenarios. However, most simulation frameworks tailored for the study of distributed systems offer no or little abstractions or models of storage resources.

In this work we extend SimGrid, a versatile toolkit for the simulation of large-scale distributed computing systems, with storage simulation capacities. We define the required abstractions and propose a new API to handle storage components and their contents in SimGrid-based simulators. Then we characterize the performance of the fundamental storage component that are disks and derive models of these resources. Finally we list several concrete use cases of storage simulations in clusters, grids, clouds, and data centers for which the proposed extension would be beneficial.

6.3. MapReduce Computations on Hybrid Distributed Computations Infrastructures

Participants: Gilles Fedak, Julio Anjos, Asma Ben Cheikh Ahmed.

In this section we report on our efforts to provide MapReduce Computing environments on Hybrid infrastructures, i.e composed of Desktop Grids and Cloud computing environments.

6.3.1. BIGhybrid - A Toolkit for Simulating MapReduce in Hybrid Infrastructures

Cloud computing has increasingly been used as a platform for running large business and data processing applications. Although clouds have become extremely popular, when it comes to data processing, their use incurs high costs. Conversely, Desktop Grids, have been used in a wide range of projects, and are able to take advantage of the large number of resources provided by volunteers, free of charge. Merging cloud computing and desktop grids into a hybrid infrastructure can provide a feasible low-cost solution for big data analysis. Although frameworks like MapReduce have been devised to exploit commodity hardware, their use in a hybrid infrastructure raise some challenges due to their large resource heterogeneity and high churn rate. This study introduces BIGhybrid, a toolkit that is used to simulate MapReduce in hybrid environments. Its main goal is to provide a framework for developers and system designers that can enable them to address the issues of Hybrid MapReduce. In this paper, we describe the framework which simulates the assembly of two existing middleware: BitDew- MapReduce for Desktop Grids and Hadoop-BlobSeer for Cloud Computing. The experimental results that are included in this work demonstrate the feasibility of our approach.

6.3.2. Parallel Data Processing in Dynamic Hybrid Computing Environment Using MapReduce

In this work, we propose a novel MapReduce computation model in hybrid computing environment called HybridMR is proposed. Using this model, high performance cluster nodes and heterogeneous desktop PCs in Internet or Intranet can be integrated to form a hybrid computing environment. In this way, the computation and storage capability of large-scale desktop PCs can be fully utilized to process large-scale datasets. HybridMR relies on a hybrid distributed file system called HybridDFS, and a time-out method has been used in HybridDFS to prevent volatility of desktop PCs, and file replication mechanism is used to realize reliable storage. A new node priority-based fair scheduling (NPBFS) algorithm has been developed in HybridMR to achieve both data storage balance and job assignment balance by assigning each node a priority through quantifying CPU speed, memory size and I/O bandwidth. Performance evaluation results show that the proposed hybrid computation model not only achieves reliable MapReduce computation, reduces task response time and improves the performance of MapReduce, but also reduces the computation cost and achieves a greener computing mode.

6.3.3. Ensuring Privacy for MapReduce on Hybrid Clouds Using Information Dispersal Algorithm

MapReduce is a powerful model for parallel data processing. The motivation of this work is to allow running map-reduce jobs partially on untrusted infrastructures, such as public Clouds and Desktop Grid, while using a trusted infrastructure, such as private cloud, to ensure that no outsider could get the 'entire' information. Our idea is to break data into meaningless chunks and spread them on a combination of public and private clouds so that the compromise would not allow the attacker to reconstruct the whole data-set. To realize this, we use the Information Dispersion Algorithms (IDA), which allows to split a file into pieces so that, by carefully dispersing the pieces, there is no method for a single node to reconstruct the data if it cannot collaborate with other nodes. We propose a protocol that allows MapReduce computing nodes to exchange the data and perform IDA-aware MapReduce computation. We conduct experiments on the Grid'5000 testbed and report on performance evaluation of the prototype.

6.4. Using Active Data to Provide Smart Data Surveillance to E-Science Users

Participants: Gilles Fedak, Anthony Simonet.

Large scientific experiments drive scientists to use many storage and computing platforms as well as different applications, tools and analysis scripts. The resulting heterogeneous environments make data management operations challenging; the significant number of events and the absence of data integration makes it difficult to track data provenance, manage sophisticated analysis processes, and recover from unexpected situations. Current approaches often require costly human intervention and are inherently error prone. The difficulty managing and manipulating such large and highly distributed datasets also limits automated sharing and collaboration. In this collaboration with Kyle Chard and Ian Foster from Argonne National Lab and University of Chicago, we study a real world e-Science application involving terabytes of data, using three different analysis and storage platforms, and a number of applications and analysis processes. We demon- strate that using a specialized data life cycle and programming model—Active Data—we can easily implement global progress monitoring, sharing and recovery from unexpected events in heterogeneous environments and automate human tasks.

6.5. HPC Component Model

Participants: Hélène Coullon, Vincent Lanore, Christian Perez, Jérôme Richard.

6.5.1. 3D FFT and L^2C

We have studied the relevance of dealing with 3D FFT parallel algorithms with the software component model L^2C [31]. We have implemented several existing 3D FFT algorithms, and we have evaluated their performance, their scalability, and their reuse rate. Experiments made on clusters of Grid'5000 and on the Curie supercomputer up to 8192 cores show that L2C based 3D assemblies are scalable and have the same kind of performance than existing 3D libraries such as FFTW or 2DECOMP. This work confirms than components can be used for optimized HPC applications

6.5.2. Stencil Skeletons in L^2C

Mesh-based scientific simulation is an important class of scientific application which could benefit from component models. Therefore, we have studied and designed a first adaptation of the SIPSim model [33] (Structured Implicit Parallelism for scientific Simulations) to handle HPC component models. The heat equation application has been implemented on top of L^2C following this adapted SIPSim model. First experiments on clusters of Grid'5000 and on the Curie supercomputer show promising results, of which a complete analysis is still ongoing. This work is a first step toward a complete implicit parallelism stencil skeleton using L^2C.

6.5.3. Reconfigurable HPC component model

High-performance applications whose structure changes dynamically during execution are extremely complex and costly to develop, maintain and adapt to new hardware. Such applications would greatly benefit from easy reuse and separation of concerns which are typical advantages of component models. Unfortunately, no existing component model is both HPC-ready (in terms of scalability and overhead) and able to easily handle dynamic reconfiguration.

We aim at addressing performance, scalability and programmability by separating locking and synchronization concerns from reconfiguration code. To this end, we have defined *directMOD*, a component model which provides on one hand a flexible mechanism to lock subassemblies with a very small overhead and high scalability, and on the other hand a set of well-defined mechanisms to easily plug various independently-written reconfiguration components to lockable subassemblies. We evaluate both the model itself and a C++/MPI implementation called *directL2C* based on L^2C.

6.6. Security for Virtualization and Clouds

Participants: Eddy Caron, Arnaud Lefray, Jonathan Rouzaud-Cornabas.

Our framework Security Aware Models for Clouds has two purposes. The first one is, for a client, to model an IaaS application composed of virtual machines, applications, datas and communications and specify the associated security requirements. The whole modelization is contained into a XML file. The second one is the scheduling. It takes as inputs application models (XML) and the infrastructure of the cloud (currently in XML) i.e. a hierarchical set of physical machines. The scheduler encapsulates applications into virtual machines when needed and then maps virtual machines onto physical machines. The result of this scheduling is a file with the mapping i.e. a list of (VM, PM) couples.

The scheduler, as a standalone engine, can be used as simulator. But it can be interfaced with a Cloud stack (e.g. OpenStack, OpenNebula) to act as a production scheduler. This inferfacing is achieved by dynamically inferring the infrastructure model from the Cloud database and applying the decision i.e the output mapping list. Furthermore, the security policies (as input) are splitted for local security enforcement on each physical machine.

Sam4C (Security-Aware Models For Clouds) is a twofold framework, namely Sam4C-Modeler and Sam4C-Scheduler. The first is dedicated to modeling an application with the tenant's virtual machines and network interconnection. The second is is a security-aware scheduler, meaning it overrides the basic default scheduler with mainly the following enhanced capabilities

We have designed a scheduling module called SPS. This module is designed to support all the operations concerning the Cloud. It is based on the OpenStack and extends OpenStack with security aspects to fulfil the requirements of Seed4C.

6.7. Locality-aware Cooperation for VM Scheduling in Distributed Clouds

Participant: Frédéric Desprez.

In collaboration with the Ascola team (A. Lèbre, J. Pastor), ASAP team (Marin Bertier), and the Myriads team (C. Tedeschi), we worked on the design of a distributed Cloud Computing infrastructure [23]. The promotion of such infrastructures as the next platform to deliver the Utility Computing paradigm, leads to new virtual machines (VMs) scheduling algorithms leveraging peer-to-peer approaches. Although these proposals considerably improve the scalability, leading to the management of hundreds of thousands of VMs over thousands of physical machines (PMs), they do not consider the network overhead introduced by multi-site infrastructures. This overhead can have a dramatic impact on the performance if there is no mechanism favoring intra-site versus inter-site manipulations.

In 2014, we designed a new building block designed on top of a network with Vivaldi coordinates maximizing the locality criterion (*i.e.,* efficient collaborations between PMs) [12]. We combined such a mechanism with DVMS, a large-scale virtual machine scheduler and showed its benefit by discussing several experiments performed on four distinct sites of the Grid'5000 testbed. With our proposal and without changing the scheduling decision algorithm, the number of inter-site operations has been reduced by 72%. This result provides a glimpse of the promising future of using locality properties to improve the performance of massive distributed Cloud platforms.

AVIZ Project-Team

6. New Results

6.1. Highlights of the Year

We had a number of highlights this year:

- Jean-Daniel Fekete was General Chair of the IEEE VIS 2014 conference, organized for the first time ever outside of the USA, in Paris, with a record attendance.
- Aviz presented 7 articles at the IEEE VIS 2014 conference, and co-organized 3 workshops.
- Five PhD students defended this year.
- Benjamin Bach was awarded the second price in the IEEE VGTC Doctoral Dissertation Competition for his thesis "Connections, Changes, Cubes: Unfolding Dynamic Networks for Visual Exploration" [10].
- Yvonne Jansen was awarded the second price for the Gilles Kahn dissertation award for her thesis "Physical and Tangible Information Visualization" [11].
- Samuel Huron received the best paper honorable mention award at DIS 2014 for the paper “Constructive Visualization” [28].

6.2. Effectiveness of Staggered Animations

Participants: Fanny Chevalier, Pierre Dragicevic [correspondant], Steven Franconeri.

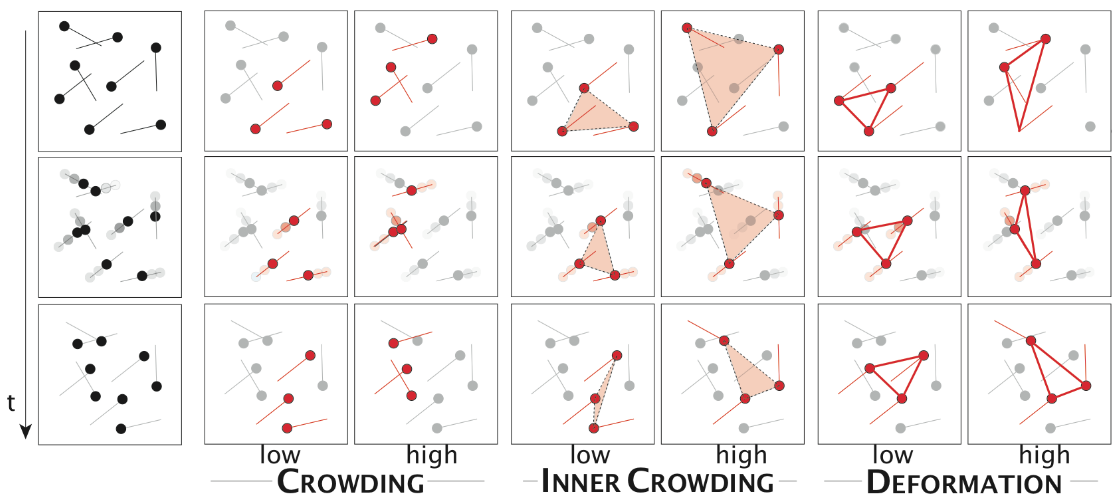


Figure 7. Illustration of the complexity metrics used in the study.

Interactive visual applications often rely on animation to transition from one display state to another. There are multiple animation techniques to choose from, and it is not always clear which should produce the best visual correspondences between display elements. One major factor is whether the animation relies on staggering—an incremental delay in start times across the moving elements. It has been suggested that staggering may reduce occlusion, while also reducing display complexity and producing less overwhelming animations, though no empirical evidence has demonstrated these advantages. Work in perceptual psychology does show that reducing occlusion, and reducing inter-object proximity (crowding) more generally, improves performance in multiple object tracking.

We empirically investigated the effectiveness of staggering [15]. We ran simulations confirming that staggering can in some cases reduce crowding in animated transitions involving dot clouds (as found in, e.g., animated 2D scatterplots). We empirically evaluated the effect of two staggering techniques on tracking tasks, focusing on cases that should most favour staggering. We found that introducing staggering has a negligible, or even negative, impact on multiple object tracking performance. The potential benefits of staggering may be outweighed by strong costs: a loss of common-motion grouping information about which objects travel in similar paths, and less predictability about when any specific object would begin to move. Staggering may be beneficial in some conditions, but they have yet to be demonstrated. Our results are a significant step toward a better understanding of animation pacing, and provide direction for further research.

More on the project Web page: fannychevalier.net/animations

6.3. Tablet-Based Interaction for Immersive 3D Data Exploration

Participants: David López, Lora Oehlberg, Candemir Doger, Tobias Isenberg [correspondant].

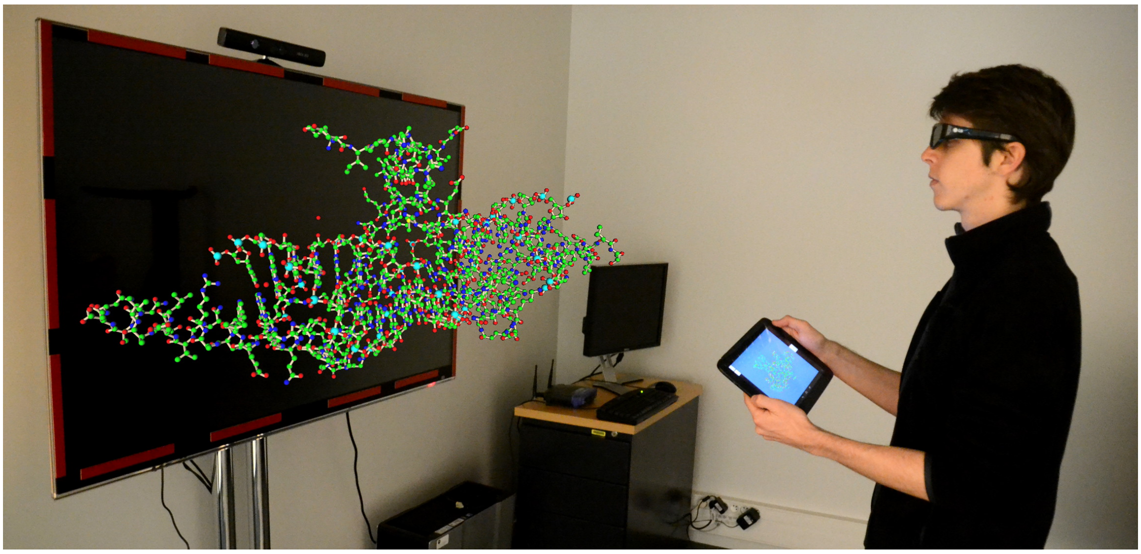


Figure 8. Illustration of the interaction setup for a combined touch-based navigation and stereoscipic viewing of 3D data.

We examined touch-based navigation of 3D visualizations in a combined monoscopic and stereoscopic viewing environment [32] (see Figure 8). We identified a set of interaction modes, and a workflow that helps users transition between these modes to improve their interaction experience. For this purpose we analyzed, in particular, the control-display space mapping between the different reference frames of the stereoscopic and monoscopic displays. We showed how this mapping supports interactive data exploration, but may also lead to conflicts between the stereoscopic and monoscopic views due to users' movement in space; we resolved these problems through synchronization. To support our discussion, we conducted an exploratory observational evaluation with domain experts in fluid mechanics and structural biology. These experts explored domain-specific datasets using variations of a system that embodies the interaction modes and workflows; we could report on their interactions and qualitative feedback on the system and its workflow.

6.4. Understanding the Perception of Star Glyphs

Participants: Johannes Fuchs, Petra Isenberg [correspondant], Anastasia Bezerianos, Fabian Fischer, Enrico Bertini.

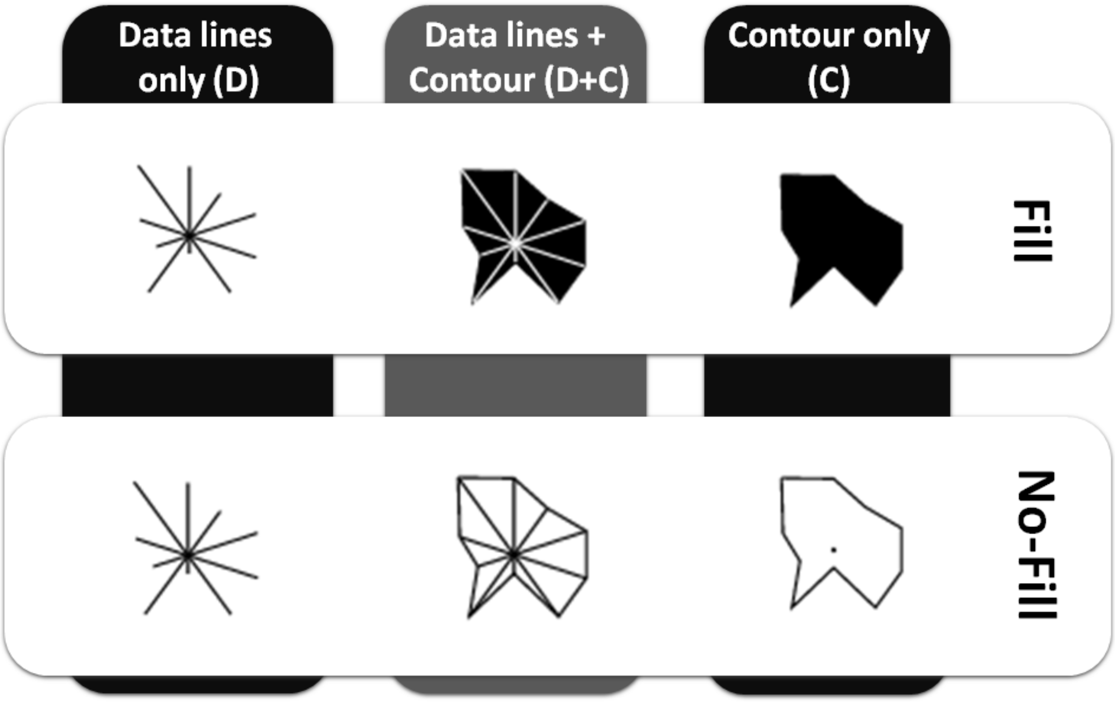


Figure 9. Illustration of the design space of the perception study.

We conducted three experiments to investigate the effects of contours on the detection of data similarity with star glyph variations [16]. A star glyph is a small, compact, data graphic that represents a multi-dimensional data point. Star glyphs are often used in small-multiple settings, to represent data points in tables, on maps, or as overlays on other types of data graphics. In these settings, an important task is the visual comparison of the data points encoded in the star glyph, for example to find other similar data points or outliers. We hypothesized that for data comparisons, the overall shape of a star glyph—enhanced through contour lines—would aid the viewer in making accurate similarity judgments. To test this hypothesis, we conducted three experiments. In our first experiment, we explored how the use of contours influenced how visualization experts and trained novices chose glyphs with similar data values. Our results showed that glyphs without contours make the detection of data similarity easier. Given these results, we conducted a second study to understand intuitive notions of similarity. Star glyphs without contours most intuitively supported the detection of data similarity. In a third experiment, we tested the effect of star glyph reference structures (i.e., tickmarks and gridlines) on the detection of similarity. Surprisingly, our results show that adding reference structures does improve the correctness of similarity judgments for star glyphs with contours, but not for the standard star glyph. As a result of these experiments, we conclude that the simple star glyph without contours performs best under several criteria, reinforcing its practice and popularity in the literature. Contours seem to enhance the detection of other types of similarity, e.g., shape similarity and are distracting when data similarity has to be judged. Based on these findings we provide design considerations regarding the use of contours and reference structures on star glyphs.

6.5. Constructive Visualization

Participants: Samuel Huron [correspondant], Yvonne Jansen, Sheelagh Carpendale.

Figure 10. Constructing a visualization with tokens: right hand positions tokens, left hand points to the corresponding data.

The accessibility of infovis authoring tools to a wide audience has been identified as one of the major research challenges. A key task of the authoring process is the development of visual mappings. While the infovis community has long been deeply interested in finding effective visual mappings, comparatively little attention has been placed on how people construct visual mappings. We conducted a study designed to shed light on how people spontaneously transform data into visual representations [18]. We asked people to create, update and explain their own information visualizations using simple materials such as tangible building blocks. We learned that all participants, most of whom had no experience in visualization, were readily able to create and talk about their own visualizations. On the basis of our observations, we discussed the actions of our participants in the context of the development of their visual representations and their analytic activities. From this we suggested implications for tool design that can enable broader support for infovis authoring.

More on the project Web page: constructive.gforge.inria.fr

6.6. Multi-touch Gestures for Data Graphics

Participants: Wesley Willett, Qi Lan, Petra Isenberg.

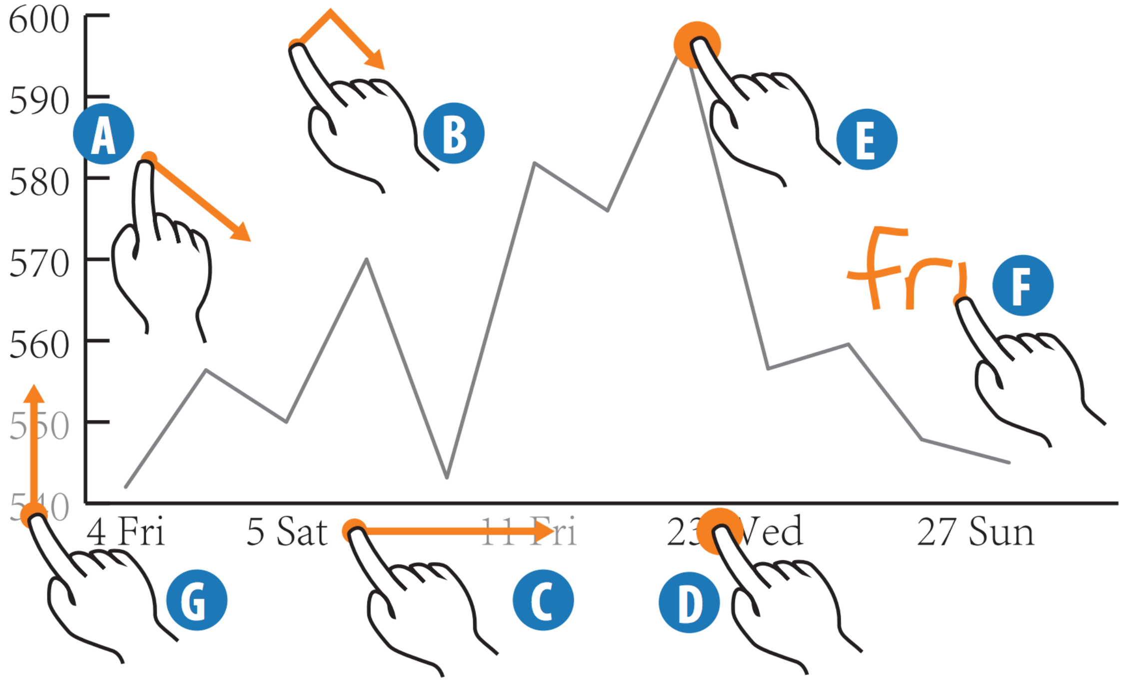


Figure 11. The most common gestures used for selecting (a) downward trends, (b) peaks, (c) ordinal ranges, (d) non-contiguous items, (e) highest points, (f) repeating dates, and (g) the lowest three points in a time series chart.

Selecting data items is a common and extremely important form of interaction with data graphics, and serves as the basis for many other data interaction techniques. However, interactive charting tools for multi-touch displays typically only provide dedicated multi-touch gestures for single-point selection or zooming. We conducted a study in which we used gesture elicitation to explore a wider range of possible selection interactions for multi-touch data graphics [35]. The results show a strong preference for simple, one-handed selection gestures. They also show that users tend to interact with chart axes and make figurative selection gestures outside the chart, rather than interact with the visual marks themselves. Finally, we found strong consensus around several unique selection gestures related to visual chart features.

6.7. Exploring Word-Scale Visualizations

Participants: Pascal Goffin, Wesley Willett, Jean-Daniel Fekete, Petra Isenberg.

Figure 12. Four examples of the integration of word-scale visualizations into HTML documents

We presented an exploration and a design space that characterize the usage and placement of word-scale visualizations within text documents [17]. Word-scale visualizations are a more general version of sparklines—small, word-sized data graphics that allow meta-information to be visually presented in-line with document text. In accordance with Edward Tufte's definition, sparklines are traditionally placed directly before or after words in the text. We described alternative placements that permit a wider range of word-scale graphics and more flexible integration with text layouts. These alternative placements include positioning visualizations between lines, within additional vertical and horizontal space in the document, and as interactive overlays on top of the text. Each strategy changes the dimensions of the space available to display the visualizations, as well as the degree to which the text must be adjusted or reflowed to accommodate them. We provided an illustrated design space of placement options for word-scale visualizations and identify six important variables that control the placement of the graphics and the level of disruption of the source text. We also contributed a quantitative analysis that highlights the effect of different placements on readability and text disruption. Finally, we used this analysis to propose guidelines to support the design and placement of word-scale visualizations.

More on the project Web page: www.aviz.fr/sparklificator

6.8. Assessing Visualization Literacy

Participants: Jeremy Boy, Ronald A. Rensink, Enrico Bertini, Jean-Daniel Fekete.

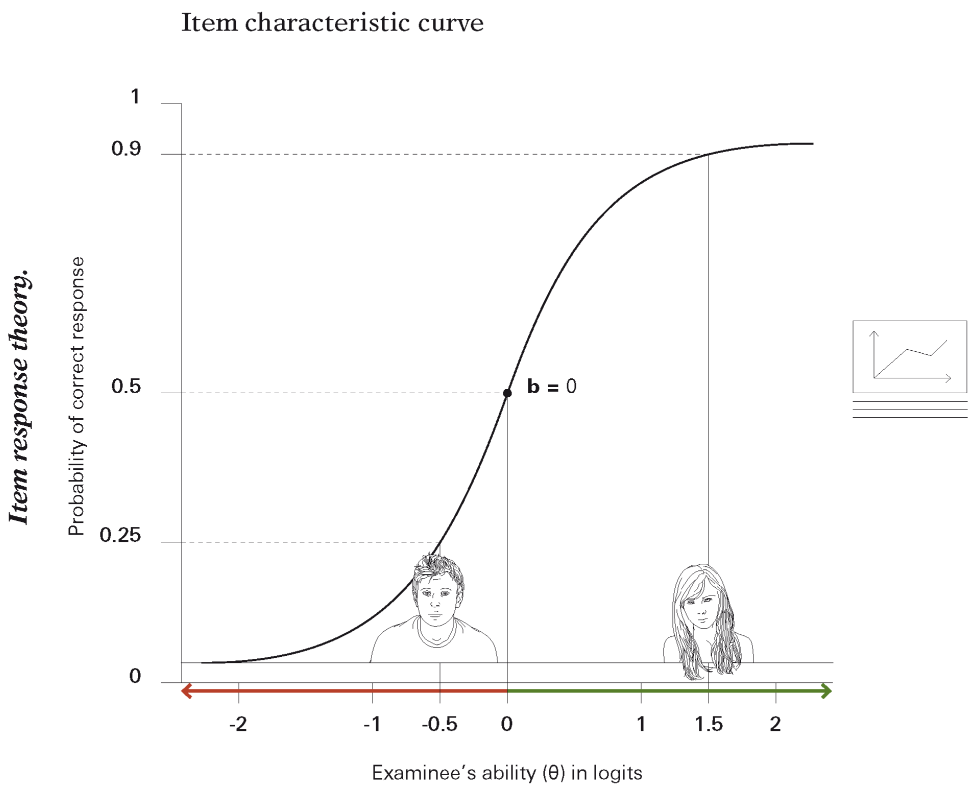


Figure 13. Example of an Item Characteristic Curve, and how people's abilities are plotted against a test-item's difficulty to determine probability of success.

We described a method for assessing the visualization literacy (VL) of a user [14]. Assessing how well people understand visualizations has great value for research (e.g., to avoid confounds), for design (e.g., to best determine the capabilities of an audience), for teaching (e. g., to assess the level of new students), and for recruiting (e. g., to assess the level of interviewees). In this project we proposed a method for assessing VL based on Item Response Theory. We described the design and evaluation of two VL tests for line graphs, and presents the extension of the method to bar charts and scatterplots. Finally, we discussed the reimplementation of these tests for fast, effective, and scalable web-based use.

More on the project Web page: peopleviz/vLiteracy/home.

AYIN Team

5. New Results

5.1. Highlights of the Year

- Yuliya Tarabalka was nominated CR1 since 1 January 2015.
- Josiane Zerubia was elected for a duration of 6 years at the board of directors of the French Society of Photogrammetry and Remote Sensing (SFPT, http://www.sfpt.fr/).
- Josiane Zerubia was invited by Technion to give a plenary talk at SIMA'14 in Ein Gedi, Israel organized for the 60th birthday of Prof. Alfred Bruckstein in May, http://www.cs.technion.ac.il/SIMA14/.

5.2. Markov Random Fields

5.2.1. Fusion of multitemporal and multiresolution remote sensing data and application to natural disasters

Participants: Ihsen Hedhli, Josiane Zerubia [contact].

This work was carried out in collaboration with Prof. Gabriele Moser and Prof. Sebastiano Serpico from DITEN departement, University of Genoa, Italy.

Multitemporal data, Multiresolution data, Supervised classification, Hierarchical Markov random fields.

The capabilities to monitor the Earth surface, and especially urban and built-up areas, from environmental disasters such as floods or earthquakes, and to assess the ground impact and damage of such events play primary roles from multiple social, economic, and human viewpoints. In this framework, accurate and time-efficient classification methods are especially important tools to support rapid and reliable assessment of the ground changes and damages induced by a disaster, in particular when an extensive area has been affected. Given the huge amount and variety of data available currently from last-generation very-high resolution (VHR) satellite missions, (such as Pléiades, COSMO-SkyMed, or WorldView-2), the main difficulty is to develop a classifier that can take benefit of multiband, multiresolution, multidate, and possibly multisensor input imagery. In such a context, Markov random field (MRF) models are widely used to solve classification problems as they permit one to integrate contextual information into the classification scheme. Due to their non-causal nature, these models generally lead to iterative inference algorithms that are computationally demanding (e.g., optimization via simulated annealing), thereby justifying the choice of a hierarchical structure, with good methodological and application-oriented properties such as: (i) the causality in scale, under Markovian assumption, which allows the use of a non-iterative algorithm with acceptable computational time and (ii) the possibility to incorporate images acquired at multiple resolutions in the hierarchy for multiresolution and multisensor fusion purposes [10]. In the proposed method, multidate and multiresolution fusion is based on explicit statistical modeling through hierarchical Markov random field modeling. The model allows both input data collected at multiple resolutions and additional multiscale features derived through wavelets to be fused. The proposed approach consists of a supervised Bayesian classifier that combines: (i) a joint class-conditional statistical model for pixelwise information and (ii) a hierarchical MRF for spatio-temporal and multiresolution contextual information. Step (i) deals, first, with the modeling of the marginal statistics of the spectral channels acquired at each resolution and conditioned to each class. Step (ii) consists in the integration of this statistical modeling in a hierarchical Markov random field for each date. An especially novel element of the proposed approach is the use of multiple quad-trees in cascade (see Figure 1), each associated with each new available image at different dates, with the aim to characterize the temporal correlations associated with distinct images in the input time series and to support the joint analysis of multitemporal, multiresolution, and possibly multisensor imagery. The transition probabilities between scales and between different dates determine the hierarchical MRF since they formalize the causality of the statistical interactions involved [11].

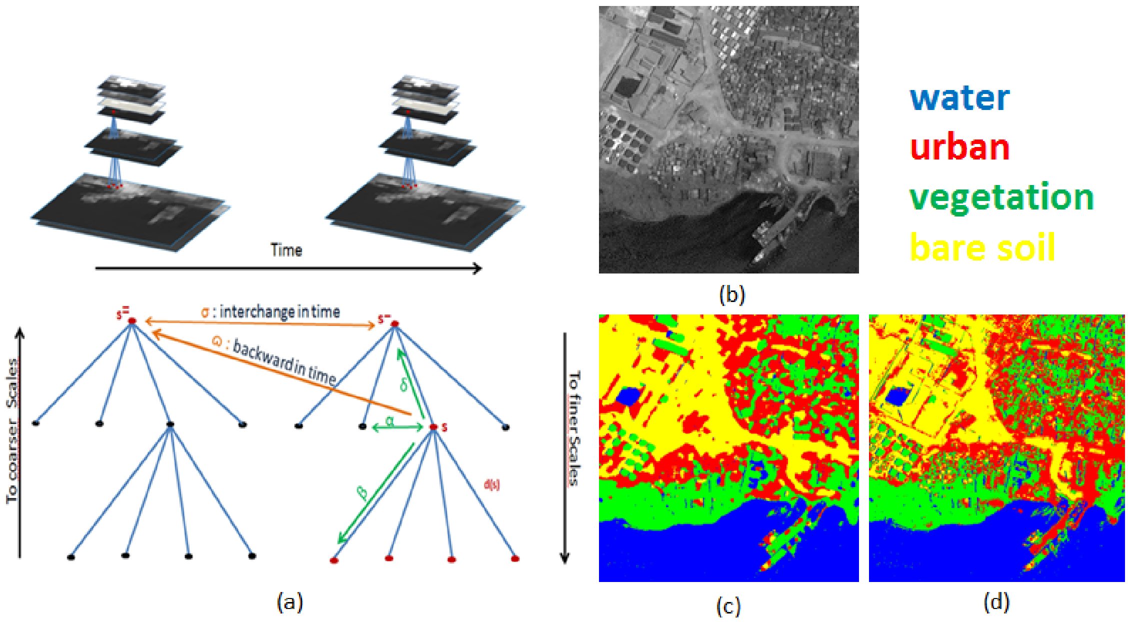


Figure 1. a) Multitemporal hierarchical structure; b) Panchromatic image of Port au Prince (Pléiades, ©CNES distribution Airbus DS, 2013); c) Classification map using single date hierarchical structure; d) Classification map obtained through the proposed multitemporal method.

5.2.2. A multi-layer Markov model for change detection in temporally separated aerial image pairs

Participant: Josiane Zerubia [contact].

This work was carried out in collaboration with Prof. Zoltan Kato from Institute of Informatics, University of Szeged, Hungary [http://www.inf.u-szeged.hu/~kato/], and Praveer Singh from Institut Mines-Telecom.

Multilayer Markov Random Fields (MRF), Histogram of Gradients (HOG), change detection, graph-cut optimization, aerial/satellite images.

In the proposed approach developed last year, we have tried to include both texture as well as pixel level information to build a three layer Markov model using the *Histogram of Oriented Gradients (HOG)* and the *Gray Level Difference* features on the topmost and bottommost layer respectively. Using a ground truth (GT) mask defined manually by an expert for each of the image pairs in the data set (obtained from the Hungarian Institute of Geodesy, Cartography and Remote Sensing), we employ a supervised technique to mark the initial set of pixels / sites as foreground or background. On the basis of the *HOG difference* and the *Gray level difference* feature vector corresponding to all the pixels in the image pair, a probability density function is fitted individually for the binary label set comprising of foreground and background labels using the GT. The probabilistic estimate is calculated using one training image pair for each data set. Using this probabilistic measure, a negative log likelihood is computed for each pixel (for both the features as well as the binary label set) which is then passed to the energy function of the proposed 3-layer MRF model. The final segmentation is obtained by minimizing the energy using a graph-cut algorithm, and subsequently a final foreground and background labelling is obtained over the combined layer. Figure 2 , shows aerial image pairs, one of them captured in 1984 by FOMI, Hungary (a) and the other one by GoogleEarth in 2007 (c). (b) is the ground truth and (d) is a combination of the hierarchical MRF based change detection (in red), ground truth (in green) and changes detected correctly (in yellow). This year, we have made a comparison of this method with two other multilayer MRFs for change detection developed at MTA-SZTAKI in Budapest, Hungary.

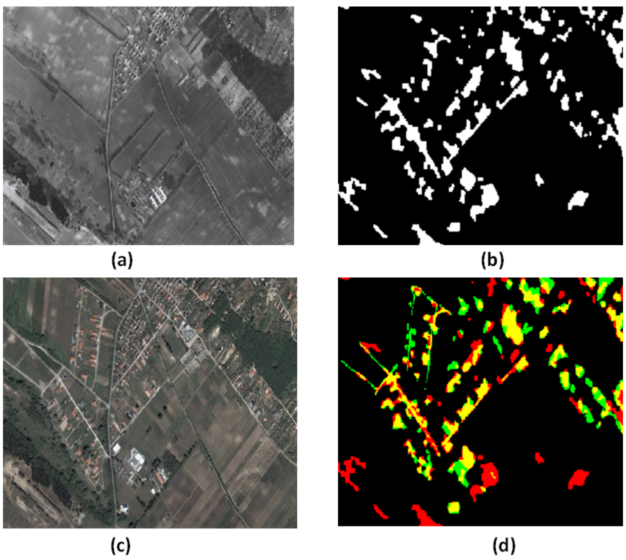


Figure 2. Change detection in an aerial image pair using a hierarchical MRF. a) Aerial image captured in 1984 by ©FOMI; b) Ground truth; c) Aerial image captured by ©GoogleEarth in 2007; d) Combination of the hierarchical MRF based change detection (in red), ground truth (in green), and changes detected correctly (in yellow).

5.2.3. Graph-cut-based model for spectral-spatial classification of hyperspectral images

Participant: Yuliya Tarabalka [contact].

This work has been done in collaboration with Aakanksha Rana (Institut Mines-Telecom/EURECOM).

Hyperspectral images, graph cut, multi-label alpha expansion, contextual information, energy minimization

The very high spatial and spectral resolution of the last generation of remote sensors provides rich information about every pixel in an image scene, hence opening new perspectives in classification, but also presenting the challenge of analysing high data volumes. While pixel-wise classification methods analyze each pixel independently, classification results can be significantly improved by including spatial information in a classifier.

In this work, we proposed a spectral-spatial method for hyperspectral image classification based on a graph cut [15]. The classification task is expressed as an energy minimization problem on the spatio-temporal graph of image pixels, and is solved by using the graph-cut α-expansion approach. The energy to optimize is computed as a sum of data and interaction energy terms, respectively. The data energy term is computed using the outputs of the probabilistic support vector machines classification. The second energy term, which expresses the interaction between spatially adjacent pixels in the eight-neighborhood, is computed by using dissimilarity measures between spectral vectors, such as vector norms, spectral angle map, or spectral information divergence. The performance of the proposed method was validated on hyperspectral images captured by the ROSIS and the AVIRIS sensors. Figure 3 compares classification results obtained by applying support vector machines and the proposed approach for the ROSIS hyperspectral image acquired over the University of Pavia. The new method yields higher classification accuracies when compared to the recent state-of-the-art approaches.

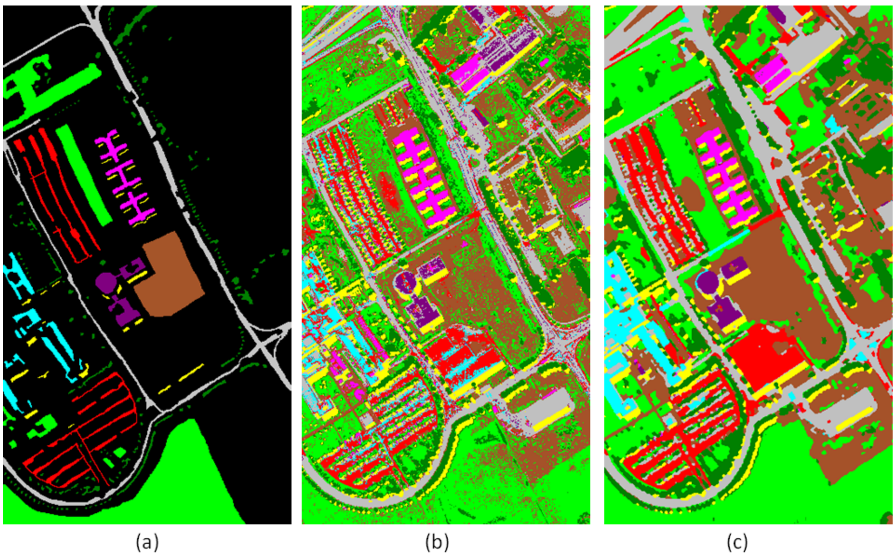


Figure 3. Hyperspectral image of the University of Pavia. (a) Ground-truth (b) Support vector machines classification map. (c) Graph-cut classification map.

5.3. Marked point processes

5.3.1. Multiple target tracking using spatio-temporal marked point processes

Participants: Paula Craciun, Josiane Zerubia [contact].

This work has been done in collaboration with Mathias Ortner from Airbus D&S (http://www.space-airbusds.com/fr/)

Multiple target tracking, stochastic geometry, point processes, remote sensing

Tracking can be defined as the problem of estimating the trajectories of objects in the image plane, as they move around the scene. Hence, a tracker assigns consistent labels to the objects in different frames of a sequence of images and can additionally provide information about the orientation, shape or size of the objects. Multi-target tracking has been historically achieved using sequential techniques, the major drawback of such methods residing in the impossibility to modify past results in the light of new data. However, applications such as offline video processing or information retrieval are not sequential in nature. Batch processing methods are preferred in this case since they do not suffer from the limitations of sequential methods. Nevertheless, these techniques remain poorly explored and highly underused.

We propose a novel approach based on spatio-temporal marked point processes to detect and track moving objects in a batch of high resolution images [17]. We develop a new, intuitive energy based model consisting of several terms that take into account both the image evidence and physical constraints such as target dynamics, track persistence and mutual exclusion. We construct a suitable optimization scheme that allows us to find strong local minima of the proposed highly non-convex energy [9]. The model has been validated on two types of data: remotely sensed satellite image sequences, characterized by high resolution, high signal to noise ratio and low temporal frequency; and biological image sequences, characterized by high resolution, low signal to noise ratio and high temporal frequency.

Tracking results are shown in Figure 4 , which shows the detection (dots) and tracking (lines) results of boats in a sequence of 14 high resolution remotely sensed images. The images are captured with a low temporal frequency at different acquisition angles.

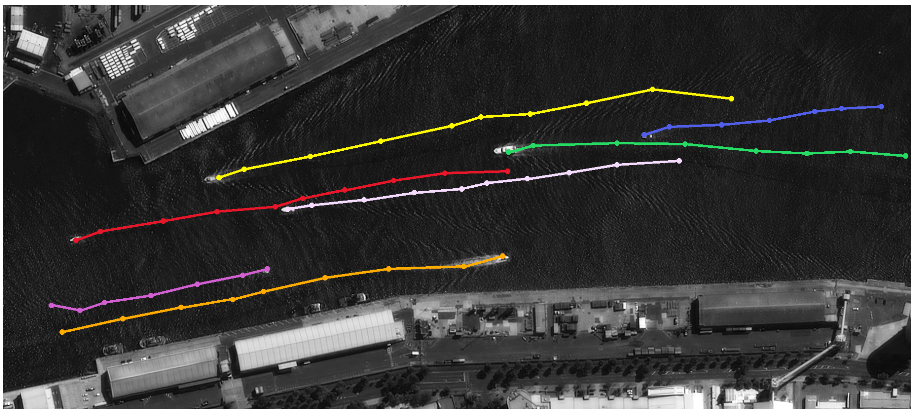


Figure 4. Detection and tracking results on a sequence of satellite images taken at different angles ©Inria/AYIN.

5.3.2. Initialization and estimation of parameters for marked point processes applied to automatic object detection on satellite images

Participants: Aurélie Boisbunon, Josiane Zerubia [contact].

This work has been done in collaboration with Rémi Flamary (Université de Nice Sophia Antipolis), Alain Rakotomamonjy (Université de Rouen) et Alain Giros (CNES). It was partially funded by the French Spatial Agency CNES [http://www.cnes.fr].

Sparse representations, large scale, stochastic algorithms, machine learning, image processing

Marked point processes (MPP) strongly rely on parameters, whose estimation affects both computation time and performances. In this work, we proposed two approaches: the first one consists in initializing MPPs with a first coarse solution obtained very quickly from sparse regularization methods [7], while the second one estimates the parameters by the Stochastic Approximation Expectation-Maximization (SAEM) algorithm [8]. We give details on both approaches below.

The first coarse solution is obtained from a deterministic sparse regularization method. This method is based on the representation of an image with objects as a sum of convolutions between atoms of a dictionary and matrices of positions of the objects in the image. Such a representation is displayed on Figure 5 . The atoms of the dictionary are fixed in advance and correspond to different instances of the objects (scales, angles, shapes, etc). This way, we transform the problem of object detection into the problem of estimating extremely sparse matrices. The algorithm we derived for solving the associated optimization problem is both parallelized and very efficient.

Up to recently, the parameters of MPPs were estimated by the Stochastic Expectation-Maximization (SEM) algorithm developed by Celeux & Diebolt (1985). This algorithm consists in alternatively estimating the expected pseudo-likelihood based on a random configuration and updating the parameter value by maximum of pseudo-likelihood. However, since it does not have a pointwise convergence, Ben Hadj et al. (2010) considered running a simulated annealing scheme after few iterations of SEM in order to reach convergence, at the cost of a higher computational time. Instead, we proposed to adapt the Stochastic Approximation Expectation-Maximization (SAEM) algorithm, developed by Delyon et al. (1999), to MPPs. Indeed, it both offers pointwise convergence and a similar computational time as SEM by efficiently taking into account past configurations in the update of the expected pseudo-likelihood.

Using both approaches resulted in the division of the computational time of the estimation of MPPs parameters by 2 and in an increase in performance of detection.

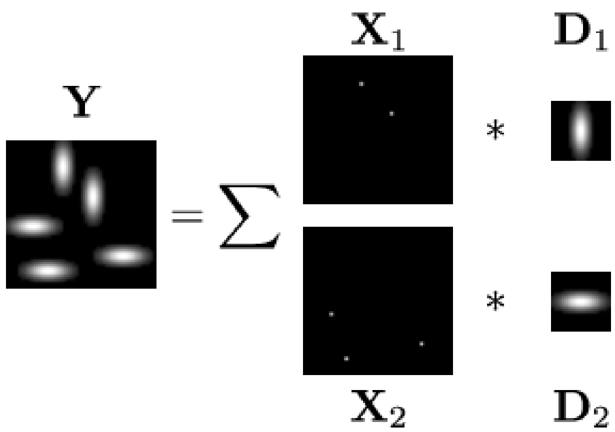


Figure 5. Representation of an image as a sum of convolutions between atoms of a dictionary and matrices of positions.

5.3.3. Generic curvilinear structure modeling via marked point process theory

Participants: Seong-Gyun Jeong, Yuliya Tarabalka, Josiane Zerubia [contact].

Curvilinear structure extraction, object detection, marked point process, stochastic inference

We proposed a marked point process model to analyze underlying curvilinear structure for wide ranges of input data, for instance, wrinkles, DNA filaments, road cracks, and blood vessels [12], [13]. It is based on sampling technique so that the model represents an arbitrary shape of the line network with a set of small line segments. The line segments should be fit into the given image data, and be harmonic with those of neighborhoods. To take these issues into consideration, we formulate a maximum a posteriori (MAP) estimation as an energy minimization problem. The energy function for given line configuration $\mathbf{s}$ can be decomposed into data likelihood term E_{data} and prior term E_{prior}:

$$E(\mathbf{s}) = \sum_{i}^{\#(\mathbf{s})} E_{\text{data}}(s_i) + \lambda \sum_{i \sim j} E_{\text{prior}}(s_i, s_j),$$

where $\#(\mathbf{s})$ denotes the total number of line segments in the current configuration, $i \sim j$ represents symmetric neighborhood system, and λ controls the relative importance of two terms. For the data term, we exploit oriented gradient information and homogeneity of the pixel intensities corresponding to line segment on the image site. The prior energy defines topology of the line configuration in that penalizes overlapping and attracts smooth connections. Another contribution of the work is to reduce parameter dependencies of the marked point process model using aggregation approach. We repeated to perform Markov chain Monte Carlo (MCMC) sampling with different parameter vectors to obtain multiple line hypotheses. Then, we combine line hypotheses to maximize the consensus among detection results.

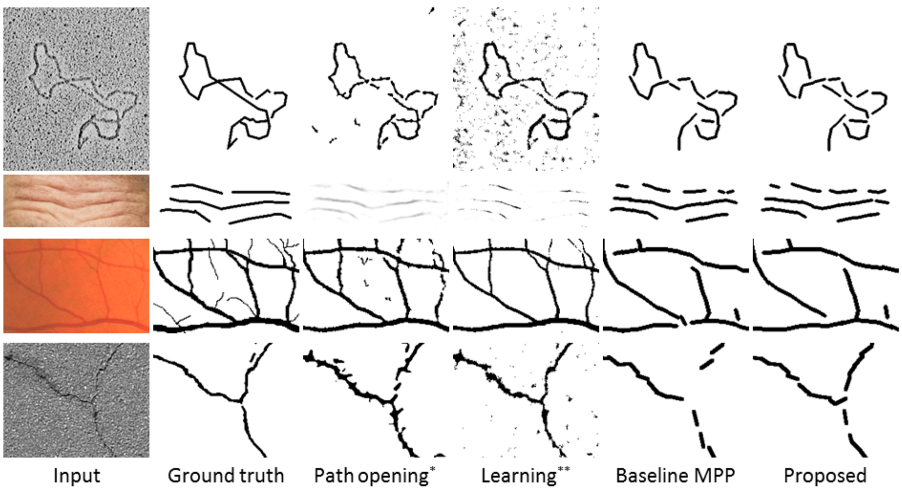


Figure 6. Comparison of the line detection results on DNA filaments, wrinkles, retina, and road cracks (top to bottom).

In figure 6 , we have compared line detection results of manually labeled image, morphological filtering (path opening), supervised feature learning, an MPP model using single parameter vector, and the proposed algorithm. The proposed algorithm extracts the most salient line structures for all datasets without any parameter estimation procedure.

5.4. Shapes and contours

5.4.1. Riemannian metrics on spaces of curves and surfaces

Participant: Ian Jermyn [contact].

This work is being done in collaboration with Anuj Srivastava of Florida State University [https://www.fsu.edu/].

Shape, Riemannian, metric, elastic, curve, surface, functional data, alignment

Statistical shape modelling has many applications in image processing and beyond. One of the key problems in this area is to develop and understand measures of shape similarity. One approach uses Riemannian metrics induced on 'shape space' by Riemannian metrics on the space of embeddings. Current work is focused on generalizing to surfaces the elastic metric used for curve embeddings, and in finding surface representations that simplify computations in the same way that the square root velocity representation simplifies computations in the case of curves. The notion of a 'square-root normal field' (SRNF), which leads to a reduced version of the full elastic metric, is a promising possibility in this direction.

The most recent work [16] has focused on estimating the inverse of the SRNF map. If this can be done even approximately, a very efficient framework results: the surfaces, represented by their SRNFs, can be efficiently analyzed using standard Euclidean tools, and only the final results need to be mapped back to the surface space. In this work, we developed a procedure for inverting SRNF maps of star-shaped surfaces, a special case for which analytic results can be obtained. We tested our method via the classification of 34 cases of ADHD (Attention Deficit Hyperactivity Disorder), plus controls, in the Detroit Fetal Alcohol and Drug Exposure Cohort study. We obtained state-of-the-art results.

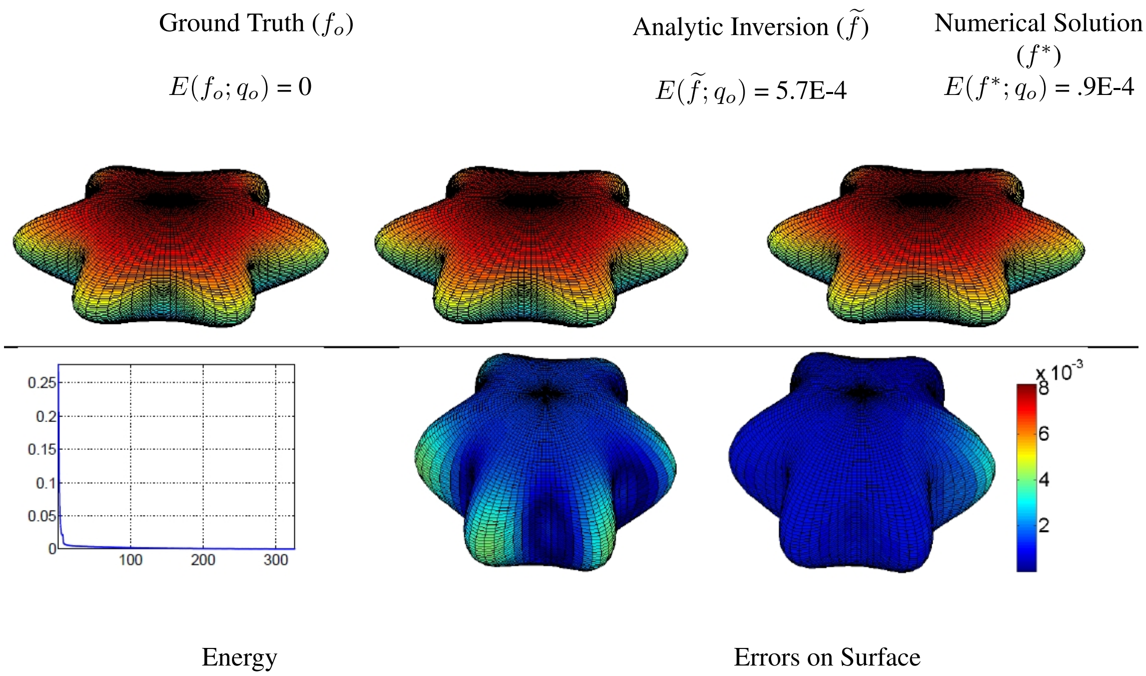


Figure 7. Reconstructing a surface from its SRNF. A target surface (f_o) is numerically reconstructed as f^* with initialization as the unit sphere. The energy plot shows the evolution of energy against iterations with initialization as a unit sphere. The analytically inverted surface $\widetilde{f}$ is shown for comparison. The corresponding energies $E(\widetilde{f}; q_o)$ and $E(f^*; q_o)$ are also shown. The errors between the reconstructed surfaces and the ground truth are shown on the ground truth surface with colours representing the magnitudes, i.e. $|f^*(s) - f_o(s)|$ for all $s \in S^2$.

5.4.2. Enforcing monotonous shape growth or shrinkage in video segmentation

Participant: Yuliya Tarabalka [contact].

This work has been done in collaboration with Dr. Guillaume Charpiat (STARS team, Inria-SAM), Dr. Bjoern Menze (Technische Universität München, Germany and Asclepios team, Inria-SAM), and Dr. Ludovic Brucker (NASA GSFC, USA) [http://www.nasa.gov].

Video segmentation, graph cut, shape analysis, shape growth

Automatic segmentation of objects from video data is a difficult task, especially when image sequences are subject to low signal-to-noise ratio or low contrast between the intensities of neighboring structures. Such challenging data are acquired routinely, for example, in medical imaging or satellite remote sensing. While individual frames can be analyzed independently, temporal coherence in image sequences provides a lot of information not available for a single image. In this work, we focused on segmenting shapes that grow or shrink monotonically in time, from sequences of extremely noisy images.

We proposed a new method for the joint segmentation of monotonically growing or shrinking shapes in a time sequence of images with low signal-to-noise ratio [3]. The task of segmenting the image time series is expressed as an optimization problem using the spatio-temporal graph of pixels, in which we are able to impose the constraint of shape growth or shrinkage by introducing unidirectional infinite-weight links connecting pixels at the same spatial locations in successive image frames. The globally-optimal solution is computed with a graph-cut algorithm. The performance of the proposed method was validated on three applications: segmentation of melting sea ice floes; of growing burned areas from time series of 2D satellite images; and of a growing brain tumor from sequences of 3D medical scans. In the latter application, we imposed an additional inter-sequences inclusion constraint by adding directed infinite-weight links between pixels of dependent image structures. Figure 8 shows a multi-year sea ice floe segmentation result. The proposed method proved to be robust to high noise and low contrast, and to cope well with missing data. Moreover, it showed linear complexity in practice.

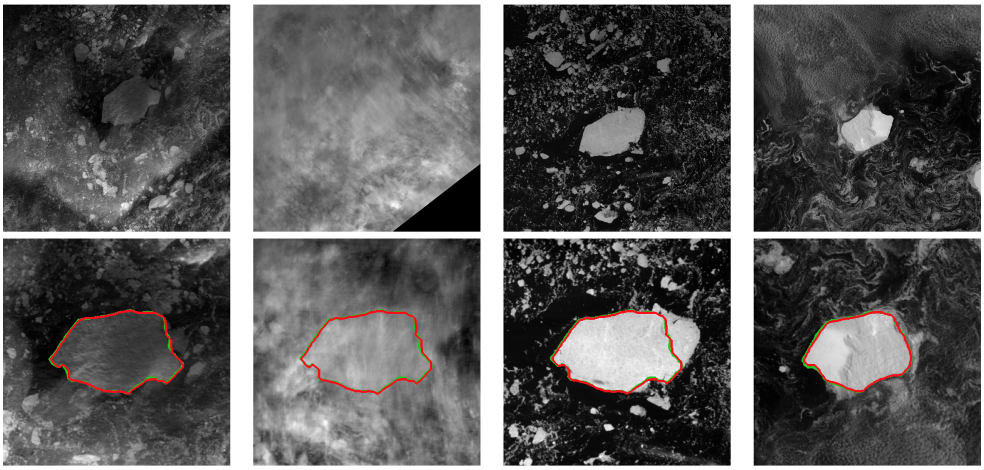


Figure 8. Top: MODIS images for four time moments (days 230, 233, 235 and 267 of 2008, respectively). Bottom: corresponding aligned images with segmentation contours (in red). Manual segmentation is shown in green.

5.4.3. Multi-label image segmentation with partition trees and shape prior

Participants: Emmanuel Maggiori, Yuliya Tarabalka [contact].

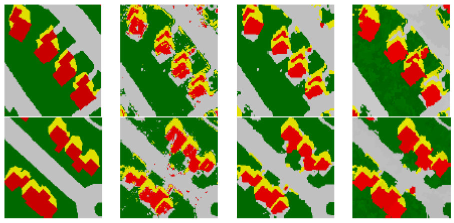
This work has been done in collaboration with Dr. Guillaume Charpiat (STARS team, Inria-SAM).

Partition trees, multi-class segmentation, shape priors, graph cut

The multi-label segmentation of images is one of the great challenges in computer vision. It consists in the simultaneous partitioning of an image into regions and the assignment of labels to each of the segments. The problem can be posed as the minimization of an energy with respect to a set of variables which can take one of multiple labels. Throughout the years, several efforts have been done in the design of algorithms that minimize such energies.

We proposed a new framework for multi-label image segmentation with shape priors using a binary partition tree [19]. In the literature, such trees are used to represent hierarchical partitions of images, and are usually computed in a bottom-up manner based on color similarities, then processed to detect objects with a known shape prior. However, not considering shape priors during the construction phase induces mistakes in the later segmentation. This paper proposes a method which uses both color distribution and shape priors to optimize the trees for image segmentation. The method consists in pruning and regrafting tree branches in order to minimize the energy of the best segmentation that can be extracted from the tree. Theoretical guarantees help reducing the search space and make the optimization efficient (see Figure 9 (i)) and [19]. Our experiments (see Figure 9) show that the optimization approach succeeds in incorporating shape information into multi-label segmentation, outperforming the state-of-the-art.

(a) Color image. (b) Manual segment. (c) BPT optimization.



(d) Manual. (e) SVM. (f) GC. (g) BPT Opt.

	$\mathcal{A}$	$\mathcal{D}$
SVM	0.764	0.50
GC	0.799	0.49
TC	0.804	0.47
Opt.	0.813	0.69

(h) Performance.

Time [s]	Time [s] (Props. 1-2)
214	84

(i) Execution time.

Figure 9. Classification results for the satellite image over Brest. $\mathcal{A}$ denotes overall classification accuracy, and $\mathcal{D}$ denotes average building's overlap. The performance of the proposed binary partition tree (BPT) optimization method is compared with the following methods: 1) support vector machines (SVM) classification; 2) graph cut (GC) with α-expansion; 3) cut on the BPT, regularized by the number of regions without using shape priors (TC).

5.5. Other detection approaches

5.5.1. Image-based evaluation of treatment responses of facial wrinkles using LDDMM registration and Gabor features

Participants: Nazre Batool, Josiane Zerubia [contact].

Face, skin texture, detection of wrinkles, LDDMM registration, response to treatment, Gabor filters, morphological processing

The goal of this work is to evaluate quantitatively the subtle variations in facial wrinkles for the same subject in response to treatment using image-based analysis. The novelty of this application is that a series of images of the same subject over a shorter time period of weeks are analyzed instead of more prevalent inter-person analysis of facial images. To overcome the challenges of detecting and evaluating such subtle changes, we propose a framework to compare image features in key wrinkle sites only while excluding the noise introduced by changes in surrounding skin texture. After initial registration using facial landmarks such as corners of eyes, nose, mouth, we propose a method based on Large Deformation Diffeomorphic Metric Mapping (LDDMM) to achieve finer registration. Fig. 10 (1a-1e) shows an example of registration using LDDMM for a pair of images.

Then we use N. Batool's previously proposed algorithm (Nazre & Chellappa (2015)) to detect key wrinkle sites. The algorithm is based on 'scaled' maximum Gabor filter responses and the incorporation of geometric constraints via morphological image processing. The binary output from the algorithm is used to create a unique wrinkle template for each subject. Fig. 10 (2a-2d) an example of obtaining a unique wrinkle template from an image using Gabor responses and wrinkle detection algorithm in (Nazre & Chellappa (2015)). Gabor responses in this template,in time series images are compared to detect subtle changes for a subject. We do not adopt the direct approach of comparing filter responses in the whole image instead of those in wrinkle template only because such an approach causes intermingling of skin texture variations in non-wrinkle sites with changes in wrinkle sites degrading the overall accuracy.

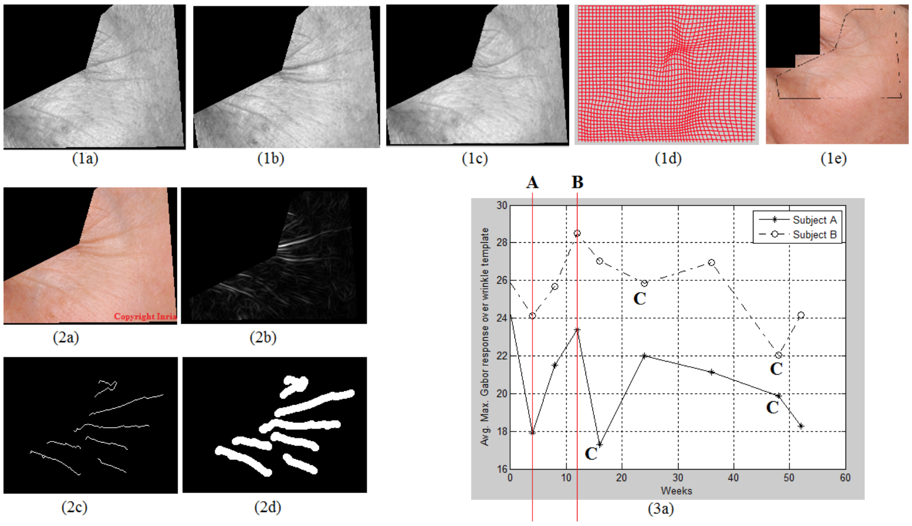


Figure 10. Overview of the evaluation framework. (1a) Week 4 image. (1b) Baseline image. (1c) Week 4 image registered using LDDMM to baseline image. (1d) Deformation of the underlying 2D space. (1e) Deformed week 4 images aligned in the original face image. (2a) Baseline image. (2b) Gabor maximum amplitude response. (2c) Detected wrinkles. (2d) The template for key wrinkle sites. (3) Plot of results for two subjects.

Fig. 10 (3a) shows a plot of results for two subjects where y-axis shows average maximum Gabor amplitude response in key wrinkle sites and x-axis corresponds to the number of weeks after the treatment. For both subjects a significant drop in the average response can been seen 4 weeks after the treatment (event 'A'). An increase in the Gabor response happened at week 12 (event 'B') which coincided with slight darkening/reddening of skin for both subjects. On the other hand, event 'C' represents co-occurrence of skin lightening with a decrease in Gabor response. These preliminary results indicate trends in wrinkle responses to treatment, skin darkening and lightening. In future, these trends will be validated by more rigorous experiments.

5.5.2. SAR data classification using generalized Gamma mixture model

Participant: Josiane Zerubia [contact].

This work has been performed in collaboration with Dr. Vladimir Krylov (University of Genoa, Italy), Prof. Heng-Chao Li, Prof. Ping-Zhi Fan (Southwest Jiaotong University, Chengdu, China) and Prof. William Emery (University of Colorado, Boulder, USA).

SAR images, statistical modeling, generalized Gamma mixture model

The accurate statistical modeling of synthetic aperture radar (SAR) images is a crucial problem in the context of effective SAR image processing, interpretation and application. In this work a semi-parametric approach is designed within the framework of finite mixture models based on the generalized Gamma distribution (GΓD) in view of its flexibility and compact analytical form. Specifically, we have developed a generalized Gamma mixture model (GΓMM) to implement an effective statistical analysis of high-resolution SAR images and proved the identifiability of such mixtures. A low-complexity unsupervised estimation method has been derived by combining the proposed histogram-based expectation-conditional maximization algorithm and the Figueiredo-Jain mixture estimation algorithm. This resulted in a numerical maximum likelihood (ML) estimator that can simultaneously determine the ML estimates of component parameters and the optimal number of mixture components. The state-of-the-art performance of the proposed method has been validated experimentally on a wide range of high-resolution SAR amplitude and intensity images.

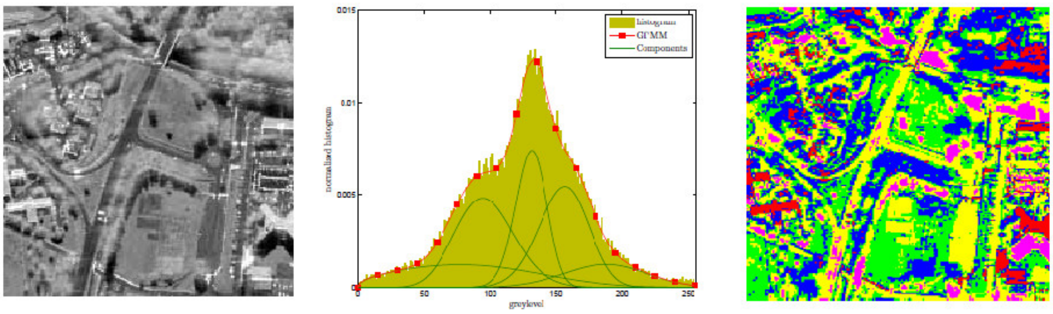


Figure 11. Statistical modeling of a RAMSES (©CNES, ONERA) image (left) by generalized Gamma mixture model (middle) and its visualization by maximum likelihood classification (right).

In Fig. 11 we demonstrate a typical result of the developed statistical modeling technique on a portion of a multilook airborne RAMSES (©CNES, ONERA) sensor acquisition over Toulouse suburbs (single polarization, downsampled to approximately 2m ground resolution). The unsupervised GΓMM estimate contains five components and reports a very accurate result that outperforms the considered benchmark statistical modeling methods. In order to visualize the estimated five statistical components we also report a maximum likelihood classification map.

BACCHUS Team

5. New Results

5.1. Penalisation methods using unstructured meshes

Participants: Heloise Beaugendre [Corresponding member], Cécile Dobrzynski, Leo Nouveau, Quentin Viville.

In Computational Fluid Dynamics the interest on embedded boundary methods for Navier-Stokes equations increases because they simplify the meshing issue, the simulation of multi-physics flows and the coupling of fluid-solid interactions in situation of large motions or deformations. Nevertheless an accurate treatment of the wall boundary conditions remains an issue of these methods. In this work we develop an immersed boundary method for unstructured meshes based on a penalization technique and we use mesh adaption to improve the accuracy of the method close to the boundary. The idea is to combine the strength of mesh adaptation, that is to provide an accurate flow description especially when dealing with wall boundary conditions, to the simplicity of embedded grids techniques, that is to simplify the meshing issue and the wall boundary treatment when combined with a penalization term to enforce boundary conditions. The bodies are described using a level-set method and are embedded in an unstructured grid. Once a first numerical solution is computed mesh adaptation based on two criteria the level-set and the quality of the solution is performed. The full paper has been published in the Journal of Computational Physics in January 2014.
External contributors. This work has benefitted from the collaboration with the University of Zurich, and in particular with R. Abgrall.

5.2. Mesh adaptation by continuous deformation

Participants: Luca Arpaia, Mario Ricchiuto [Corresponding member].

As discussed in section 3.3 Meshes and scalable discrete data structures an accurate resolution of time dependent flows requires a dynamic mesh adaptation procedure which is quite complex and costly, especially when combined with parallel distributed memory implementations. To alleviate this cost, and still allow mesh adaptation for time dependent problems we have started to look into adaptation techniques which do not involve any re-meshing. In particular, we have studied methods based on continuous mesh deformation. These methods require, at each time step, the solution of a PDE for the mesh as well as for the flow variables. This year we have settled several fundamental questions related to the basic formulation of the method, and its coupling with either implicit or explicit time discretisation methods of the flow variables. Initial applications to free surface flows have been considered showing the generality and potential of our results [39].

5.3. Non-hydrostatic modelling of free surface flows

Participants: Stevan Bellec, Mathieu Colin [Corresponding member], Andrea Filippini, Maria Kazolea, Mario Ricchiuto.

This year we have made a lot of progress in the understanding of the properties of Boussinesq-type models for near shore applications. In particular, we have performed a systematic analysis of the nonlinear behaviour of these models in the surf-zone, and in particular of their shoaling properties. These properties influence fundamentally the wave breaking process, and thus the impact of the wave on coastal structures. We have clearly identified two families of physical behaviours, associated to a similar formal structure of the equations. This result has been presented in [30], [45], and the full study is currently in revision on the Coastal Engineering journal.

In parallel, we have continued the study of the implementation of wave breaking models, comparing several physical criteria for the detection of the beginning and end of the breaking process. So far, we have only tested the so-called hybrid approach in which the hyperbolic Shallow Water equations are used in breaking regions, and the energy dissipation of breaking waves is modelled by the dissipation of mathematical entropy in shock waves. The work performed complements the initial study performed by M. Kazolea in her PhD and also proposes new physical detection criteria [28], [44] (a full paper is in preparation).

Furthermore, we have began a systematic study on the existence of particular solutions (such as solitary waves for example) to the different Boussinesq-type models in view of having efficient materials to determinate the efficiency of our numerical schemes and to perform preliminary simulations.

The last important theoretical brick we added this year is the study of fully discrete asymptotic models, obtained by pre-discretizing the two-dimensional incompressible free surface Euler equations with a finite element method, and then by performing an asymptotic development (in terms of the classical nonlinearity and dispersion parameters). We have thus obtained a discrete model which, although consistent with a known continuous Boussinesq system, represents a surprisingly improved discret eversion of these equations, hardly obtainable by classical discretisation choices.

Besides the modelling effort, we have also started woking on real applications. In particular, we have worked on case studies involving harbour dynamics and river hydraulics. In the first case, M. Kazolea has performed a systematic study of the contribution of harbour resonance in the excitation of the Venetian harbor basin of Chania, during typical winder storms. Concerning river hydraulics, we have performed a parametric study of the appearance of tidal bores in estuaries, with parameters given by the tide non-linearity (amplitude), and the friction in the river. Both works will be presented at the next world congress of the International Association for Hydro-Environment Engineering.
External contributors. This work has benefitted from the collaboration with the EPOC lab in Bordeaux, and in particular with P. Bonneton.

5.4. Two-phase flow numerical simulation with real-gas effects and occurrence of rarefaction shock waves

Participants: Maria Giovanna Rodio, Pietro Marco Congedo [Corresponding member].

A discrete equation method (DEM) for the simulation of compressible multiphase flows including real-gas effects has been developed. A reduced five equation model is obtained starting from the semi-discrete numerical approximation of the two-phase model. A simple procedure is then proposed for using a more complex equation of state, thus improving the quality of the numerical prediction. Classical test-cases well-known in literature are performed featuring a strong importance of thermodynamic complexity for a good prediction of temperature evolution. Finally, a computational study on the occurrence of rarefaction shock waves (RSW) in a two-phase shock tube is presented, with dense vapors of complex organic fluids. Since previous studies have shown that a RSW is relatively weak in a single-phase (vapor) configuration, its occurrence and intensity are investigated considering the influence of the initial volume fraction, initial conditions and the thermodynamic model [11]. A transition modelling has been also introduced for considering heat and mass transfer terms. In this way, metastable states have been simulated in cavitating flows Finally, a semi-intrusive stochastic technique has been formulated for taking into account uncertainties in the simulation of metastable states.
External contributors. This work has benefitted from the collaboration with the University of Zurich, and in particular with R. Abgrall.

5.5. Formulation of stochastic methods for CFD

Participants: Gianluca Geraci, Kunkun Tang, Francesca Fusi, Pietro Marco Congedo [Corresponding member].

A novel adaptive strategy for stochastic problems has been developed, inspired from the classical Harten's framework. The proposed algorithm allows building, in a very general manner, stochastic numerical schemes starting from a whatever type of deterministic schemes and handling a large class of problems, from unsteady to discontinuous solutions. Its formulations permits to recover the same results concerning the interpolation theory of the classical multiresolution approach, but with an extension to uncertainty quantification problems. The present strategy permits to build numerical scheme with a higher accuracy with respect to other classical uncertainty quantification techniques, but with a strong reduction of the numerical cost and memory requirements. Moreover, the flexibility of the proposed approach allows to employ any kind of probability density function, even discontinuous and time varying, without introducing further complications in the algorithm. The advantages of the present strategy are demonstrated by performing several numerical problems where different forms of uncertainty distributions are taken into account, such as discontinuous and unsteady custom-defined probability density functions. In addition to algebraic and ordinary differential equations, numerical results for the challenging 1D Kraichnan–Orszag are reported in terms of accuracy and convergence. Finally, a two degree-of-freedom aeroelastic model for a subsonic case is presented. Though quite simple, the model allows recovering some physical key aspect, on the fluid/structure interaction, thanks to the quasi-steady aerodynamic approximation employed. The injection of an uncertainty is chosen in order to obtain a complete parameterization of the mass matrix. All the numerical results are compared with respect to classical Monte Carlo solution and with a non-intrusive Polynomial Chaos method [3].

Moreover, in [15], an anchored ANOVA method is proposed to decompose statistical moments. Compared to standard ANOVA with mutually orthogonal components, anchored ANOVA, with arbitrary anchor point, loses orthogonality if employing the same measure. However, an advantage consists in the considerably reduced number of deterministic solver's computations, which renders uncertainty quantification of real engineering problems much easier. Different from existing methods, covariance decomposition of output variance is used in this paper to take account of interactions between non-orthogonal components, yielding an exact variance expansion, and thus, with a suitable numerical integration method, provides a strategy that converges. This convergence is verified by studying academic tests. In particular, sensitivity problem of existing method to anchor point is analyzed via Ishigami case, and we point out covariance decomposition survives from it. Covariance-based sensitivity indices (SI) are also used, compared to variance-based SI. Furthermore, we emphasize covariance decomposition can be generalized in a straightforward way to decompose high order moments. For academic problems, results show the method converges to exact solution regarding both skewness and kurtosis. Finally, the proposed method is applied on a realistic case, i.e. estimating chemical reactions uncertainties in a hypersonic flow around a space vehicle during an atmospheric reentry.

External contributors. This work has benefitted from the collaboration with the University of Zurich, and in particular with R. Abgrall.

BAMBOO Project-Team

6. New Results

6.1. Evolution of the genomes of endosymbionts in insects: the case of Hamiltonella defensa interacting with its various partners

Insect cells host many endosymbiotic bacteria, which are in general classified according to their importance for the host: “primary” symbionts are by definition mandatory and synthesize essential nutrients for the insects that feed on poor or unbalanced food sources, while “secondary” symbionts are optional and use mutualistic strategies and/or manipulation of reproduction to invade and persist within insect populations. *Hamiltonella defensa* is a secondary endosymbiont that established two distinct associations with phloemophagous insects. In aphids, it protects the host against parasitoid attacks. Its ability to infect many host tissues, notably the hemolymph, could promote its contact with parasitoid eggs. Despite this protective phenotype, the high costs associated with its presence within the host prevent its fixation in the population. In the whitefly *Bemisia tabaci* however, this symbiont is found only in cells specialised in hosting endosymbionts, the bacteriocytes. In these cells, it cohabits with other symbiotic species, such as the primary symbiont *Portiera aleyrodidarum*, a proximity that favors potential exchanges between the two symbionts. It is fixed in populations of *B. tabaci*, which suggests an important role for the consortium, probably nutritious.

We studied the specificities of each of these systems. First, in the bacteriocytes of *B. tabaci*, we identified a partitioning of the synthetic capacities of two endosymbionts, *H. defensa* and *P. aleyrodidarum*, in addition to a potential metabolic complementation between the symbionts and their host for the synthesis of essential amino acids. We proposed a key nutritive role for *H. defensa*, which would indicate a transition to a mandatory status in relation to the host and would explain its fixation in the population.

We also focused on the genomic evolution of the genus *Hamiltonella*, by comparing the strains infecting *B. tabaci* with a strain infecting the aphids. We highlighted the specialization of the symbionts to their hosts, and found that the genomes of the endosymbionts reflected their respective ecology. The aphid strain thus possesses many virulence factors and is associated with two partners, a bacteriophage and a recombination plasmid. These systems, inactive in the symbiont of *B. tabaci*, are directly related to the protection against and arms race with parasitoids. Conversely, the presumed avirulence of whitefly endosymbionts is consistent with their nutritional phenotype and a transition to a mandatory status to the host.

Finally, we studied the phenomenon of “accelerated mutation rate” in *H. defensa*, compared to its sister species *Regiella insecticola*, which is also a clade of protective endosymbionts of aphids. After excluding the assumption that the transition to the intracellular life occurred independently in the two lineages, we tried to establish a link between these differences in terms of evolvability in the endosymbionts and of their gene contents, particularly for genes involved in ecology and DNA repair. All the results obtained have provided insight into the evolution of the species *H. defensa*, since the last ancestor to the present species, by establishing a link between bacterial phenotype and genomic evolution.

The publications related to this area of research are either submitted or in preparation (to be submitted in the first months of the year).

6.2. Cardinium cBtQ1 providing insights into genome reduction, symbiont motility, and its settlement in Bemisia tabaci

Many insects harbor inherited bacterial endosymbionts. Although some of them are not strictly essential and are considered facultative, they can be a key to host survival under specific environmental conditions, such as parasitoid attacks, climate changes, or insecticide pressures. The whitefly *Bemisia tabaci* is at the top of the list of organisms inflicting agricultural damage and outbreaks, and changes in its distribution may be associated to global warming. In partnership with the group of Andrès Moya at the ICBiBE (Institut Cavanilles de Biodiversitat i Biologia Evolutiva), the genome of *Cardinium* cBtQ1, a facultative bacterial endosymbiont of *B. tabaci*, was sequenced and analysed [23].

6.3. Mitochondrial respiration and genomic analysis provide insight into the influence of the symbiotic bacterium on host trypanosomatid oxygen consumption

Certain trypanosomatids, such as *Angomonas deanei*, co-evolve with an endosymbiotic bacterium in a mutualistic relationship that is characterised by intense metabolic exchanges. We were able to show that the symbionts were able to respire for up to 4 h after isolation from the host. Moreover, our work suggests that the symbiont influences the mitochondrial respiration of the host protozoan [5].

6.4. Telling metabolic stories to explore metabolomics data

The increasing availability of metabolomics data enables to better understand the metabolic processes involved in the immediate response of an organism to environmental changes and stress. The data usually come in the form of a list of metabolites whose concentrations significantly changed under some conditions, and are thus not easy to interpret without being able to precisely visualize how such metabolites are interconnected.

We presented a method that enables to organize the data from any metabolomics experiment into metabolic stories [18]. Each story corresponds to a possible scenario explaining the flow of matter between the metabolites of interest. These scenarios may then be ranked in different ways depending on which interpretation one wishes to emphasize for the causal link between two affected metabolites: enzyme activation, enzyme inhibition or domino effect on the concentration changes of substrates and products. Equally probable stories under any selected ranking scheme can be further grouped into a single anthology that summarizes, in a unique subnetwork, all equivalently plausible alternative stories. An anthology is simply a union of such stories. We detailed an application of the method to the response of yeast to cadmium exposure. We used this system as a proof of concept for our method, and we showed that we are able to find a story that reproduces very well the current knowledge about the yeast response to cadmium. We further showed that this response is mostly based on enzyme activation. We also provided a framework for exploring the alternative pathways or side effects this local response is expected to have in the rest of the network. We discussed several interpretations for the changes we see, and we suggested hypotheses that could in principle be experimentally tested. Noticeably, our method requires simple input data and could be used in a wide variety of applications.

6.5. MiRNA and co: Methodologically exploring the world of small RNAs

We developed a reliable, robust, and much faster method for the prediction of pre-miRNAs. With this method, we aimed mainly at two goals: efficiency and flexibility. Efficiency was made possible by means of a quadratic algorithm. Since the majority of the predictors use a cubic algorithm to verify the pre-miRNA hairpin structure, they may take too long when the input is large. Flexibility relies on two aspects, the input type and the organism clade. MIRINHO can receive as input both a genome sequence and small RNA sequencing (sRNA-seq) data of both animal and plant species. To change from one clade to another, it suffices to change the lengths of the stem-arms and of the terminal loop. Concerning the prediction of plant miRNAs, because their pre-miRNAs are longer, the methods for extracting the hairpin secondary structure are not as accurate as for shorter sequences. With MIRINHO, we also addressed this problem, which enabled to provide premiRNA secondary structures more similar to the ones in MIRBASE than the other available methods.

Mirinho served as the basis to two other issues we addressed. The first issue led to the treatment and analysis of sRNA-seq data of *Acyrthosiphon pisum*, the pea aphid. The goal was to identify the miRNAs that are expressed during the four developmental stages of this species, allowing further biological conclusions concerning the regulatory system of such an organism. For this analysis, we developed a whole pipeline, called MIRINHOPIPE, at the end of which MIRINHO was aggregated.

We then moved on to the second issue, that involved problems related to the prediction and analysis of non-coding RNAs (ncRNAs) in the bacterium *Mycoplasma hyopneumoniae*. A method, called ALVINHO, was thus developed for the prediction of targets in this bacterium, together with a pipeline for the segmentation of a numerical sequence and detection of conservation among ncRNA sequences using a k-partite graph.

We finally addressed a problem related to motifs, that is to patterns, that may be composed of one or more parts, that appear conserved in a set of sequences and may correspond to functional elements. This had already been addressed in a robust method called Smile. However, depending on the input parameters, the output may be too large to be tractable, as was realized in other works of the team. We then presented some clustering solutions to group the motifs that may correspond to a same biological element, and thus to better distinguish the biologically significant ones from noise that may be present in what often are large outputs from many motif extraction algorithms.

The publications related to this area of research are either submitted or in preparation (to be submitted in the first months of the year).

6.6. Efficient Algorithms for analysing RNA-seq Data

In the last years, we had addressed the problem of identifying and quantifying variants (alternative splicing and genomic polymorphism) in RNA-seq data when no reference genome is available, without assembling the full transcripts. Based on the fundamental idea that each variant corresponds to a recognizable pattern, a bubble, in a de Bruijn graph constructed from the RNA-seq reads, we propose a general model for all variants in such graphs. We then introduced an exact algorithm, called KISSPLICE, to extract alternative splicing events. We had showed that it enables to identify more correct events than general purpose transcriptome assemblers.

The main time bottleneck in the KISSPLICE algorithm is the bubble enumeration step. Thus, in an effort to make our method as scalable as possible, we had modified Johnson's cycle listing algorithm (Johnson (1975)) to enumerate bubbles in general directed graphs, while maintaining the same time complexity. We now proposed, using a different enumeration technique, an algorithm to list bubbles with path length constraints in weighted directed graphs [29]. For a graph with n vertices and m edges, the method we propose lists all bubbles with a given source in $O(n(m + nlogn))$ delay. Moreover, we experimentally showed that this algorithm is several orders of magnitude faster than the listing algorithm of KISSPLICE to identify bubbles corresponding to alternative splicing events.

Additionally, we showed that the same techniques used to list bubbles can be applied to one classical enumeration problem: K-shortest paths problems [29]. We considered a different parameterisation of the K-shortest paths problem: instead of bounding the number of st-paths, we bound the weight of the st-paths. We present a general scheme to list bounded length st-paths in weighted graphs that takes $O(nt(n, m))$ time per path, where $t(n, m)$ is the time for a single source shortest path computation. This algorithm uses memory linear in the size of the graphs, independent of the number of paths output. For undirected non-negatively weighted graphs, we also show an improved algorithm that lists all st-paths with bounded length in $O((m + t(n, m)))$ time per path.

The main memory bottleneck in KISSPLICE is the construction and representation of the de Bruijn graph. Thus, again with the goal to make our method as scalable as possible, we propose a new compact way to build and represent a de Bruijn graph improving over the state of the art [22]. We show both theoretically and experimentally that our approach uses 30% to 40% less memory than such state of the art, with an insignificant impact on the construction time. Our de Bruijn graph representation is general, in other words it is not restricted to the variation finding or RNA-seq context, and can be used as part of any algorithm that represents NGS data with de Bruijn graphs.

A major issue when analysing transcriptomes using short sequencing reads is to be able to deal with repeats that are longer than the reads. We proposed a first explicit model for large families of inexact repeats in the de Bruijn Graphs generated from RNA-seq data [21]. Taking advantage of this modelling, we also proposed an efficient algorithm which enumerates alternative splicing events without traversing repeat-induced subgraphs, therefore offering a first answer to one the main question left open at the end of Gustavo Sacomoto's PhD [4].

Motivated by previous work on the classical problem of listing cycles, we also studied from a more purely theoretical point of view how to list chordless cycles [28]. We thus developed an amortized $\tilde{O}(|V|)$-delay algorithm for listing chordless cycles in undirected graphs. Chordless cycles are very natural structures in undirected graphs, with an important history and distinguished role in graph theory. The best known solution to list all the C chordless cycles contained in an undirected graph $G = (V, E)$ takes $O(|E|2 + |E| \cdot C)$ time. In this paper we provide an algorithm taking $\tilde{O}(|E| + |V| \cdot C)$ time. We also show how to obtain the same complexity for listing all the P chordless st-paths in G (where C is replaced by P).

6.7. Reference-free detection of isolated SNPs

Detecting single nucleotide polymorphisms (SNPs) between genomes is becoming a routine task with next-generation sequencing. Generally, SNP detec- tion methods use a reference genome. As non-model organisms are increasingly investigated, the need for reference-free methods has been amplified. Most of the existing reference-free methods have fundamental limitations: they can only call SNPs between exactly two datasets, and / or they require a prohibitive amount of computational resources. V. Lacroix participated in the developement of a method, called DISCOSNP to detect both heterozygous and homozygous isolated SNPs from any number of read datasets, without a reference genome, and with very low memory and time footprints (billions of reads can be analyzed with a standard desktop computer) [25]. To facilitate downstream genotyping analyses, DISCOSNP ranks predictions and outputs quality and coverage per allele. Compared to finding isolated SNPs using a state-of-the-art assembly and mapping approach, DISCOSNP requires significantly less computational resources, shows similar precision / recall values, and highly ranked predictions are less likely to be false positives. An experimental validation was conducted on an arthropod species (the tick *Ixodes ricinus*) on which de novo sequencing was performed. Among the predicted SNPs that were tested, 96% were successfully genotyped and truly exhibited polymorphism.

6.8. Endothelial, epithelial, and fibroblast cells exhibit specific splicing programs independently of their tissue of origin

Alternative splicing is the main mechanism of increasing the proteome diversity coded by a limited number of genes. It is well established that different tissues or organs express different splicing variants. However, organs are composed of common major cell types, including fibroblasts, epithelial, and endothelial cells. By analysing large-scale data sets generated by The ENCODE Project Consortium and after extensive RT-PCR validation, we demonstrated that each of the three major cell types expresses a specific splicing program independently of its organ origin [17]. Furthermore, by analysing splicing factor expression across samples, publicly available splicing factor binding site data sets (CLIP-seq), and exon array data sets after splicing factor depletion, we identified several splicing factors that contribute to establishing these cell type-specific splicing programs.

6.9. Length and symmetry on the sorting by weighted inversions problem

Large-scale mutational events that occur when stretches of DNA sequence move throughout genomes are called genome rearrangement events. In bacteria, inversions are one of the most frequently observed rearrangements. In some bacterial families, inversions are biased in favour of symmetry as shown by recent research. In addition, several results suggest that short segment inversions are more frequent in the evolution of microbial genomes. Despite the fact that symmetry and length of the reversed segments seem very important, they have not been considered together in any problem in the genome rearrangement field. We defined the problem of sorting genomes (or permutations) using inversions whose costs are assigned based on their lengths and asymmetries [27]. We presented five procedures and assessed their performance on small sized permutations. The ideas presented in the paper provide insights to solve the problem and set the stage for a proper theoretical analysis.

6.10. Efficient tree reconciliation enumerator plus cophylogeny reconstruction algorithm via an Approximate Bayesian Computation

Phylogenetic tree reconciliation is the approach of choice for investigating the co-evolution of sets of organisms such as hosts and parasites. It consists in a mapping between the parasite tree and the host tree using event-based maximum parsimony. Given a cost model for the events, many optimal reconciliations are however possible. Any further biological interpretation of them must therefore take this into account, making the capacity to enumerate all optimal solutions a crucial point. Only two algorithms currently exist that attempt such enumeration; in one case not all possible solutions are produced while in the other not all cost vectors are currently handled. Our objective in addressing this problem was two-fold. The first was to fill this gap, and the second was to test whether the number of solutions generally observed can be an issue in terms of interpretation.

We presented a polynomial-delay algorithm called EUCALYPT for enumerating all optimal reconciliations [12]. We showed that in general many solutions exist. We gave an example where, for two pairs of host-parasite trees having each less than 41 leaves, the number of solutions is 5120, even when only time-feasible ones are kept. To facilitate their interpretation, those solutions were also classified in terms of how many of each event they contain. The number of different classes of solutions may thus be notably smaller than the number of solutions, yet they may remain high enough, in particular for the cases where losses have cost 0. In fact, depending on the cost vector, both numbers of solutions and of classes thereof may increase considerably (for the same instance, to respectively 4080384 and 275). To further deal with this problem, we introduced and analysed a restricted version where host-switches are allowed to happen only between species that are within some fixed distance along the host tree. This restriction allows us to reduce the number of time-feasible solutions while preserving the same optimal cost, as well as to find time-feasible solutions with a cost close to the optimal in the cases where no time-feasible solution is found.

Despite an increasingly vast literature on cophylogenetic reconstructions for studying host-parasite associations, understanding the common evolutionary history of such systems remains a problem that is far from being solved. Most algorithms for host-parasite reconciliation use an event-based model, where the events include in general (a subset of) cospeciation, duplication, loss, and host-switch. All known parsimonious event-based methods then assign a cost to each type of event in order to find a reconstruction of minimum cost. This is what we did ourselves in EUCALYPT. The main problem with this approach is that the cost of the events strongly influences the reconciliation obtained.

To deal with this problem, we developed an algorithm, called COALA, for estimating the frequency of the events based on an approximate Bayesian computation approach [8]. The benefits of this method are twofold: (1) it provides more confidence in the set of costs to be used in a reconciliation, and (2) it allows estimation of the frequency of the events in cases where the dataset consists of trees with a large number of taxa.

We evaluated our method on simulated and on biological datasets. We showed that in both cases, for the same pair of host and parasite trees, different sets of frequencies for the events lead to equally probable solutions. Moreover, often these solutions differ greatly in terms of the number of inferred events. It appears crucial to take this into account before attempting any further biological interpretation of such reconciliations. More generally, we also showed that the set of frequencies can vary widely depending on the input host and parasite trees. Indiscriminately applying a standard vector of costs may thus not be a good strategy.

6.11. Others

Other works, often experimental were also developed during 2014 and published in a number of papers [6], [7], [9], [10], [11], [13], [16], [19], [20], [24], [26].

BEAGLE Project-Team

5. New Results

5.1. Highlights of the Year

We organized two satellite workshops of international conferences:

- The Aevol tutorial during ALife 2014 (July 30 - August 2, New York) http://www.aevol.fr/alifeTutorial
- The "Computational Methods and Modeling of Astrocyte Physiology and Neuron-Glia Interactions" workshop during the Computational NeuroScience 2014 conference (July 26 - 31, Quebec City, Canada)

These highlight our active presence in the scientific life of our two sub-domains in major conferences.

5.2. Sparse short-distance connections enhance calcium wave propagation in a 3D model of astrocyte networks

Participants: H. Berry, J. Lallouette, M. De Pittá

Traditionally, astrocytes have been considered to couple via gap-junctions into a syncytium with only rudimentary spatial organization. However, this view is challenged by growing experimental evidence that astrocytes organize as a proper gap-junction mediated network with more complex region-dependent properties. On the other hand, the propagation range of intercellular calcium waves (ICW) within astrocyte populations is as well highly variable, depending on the brain region considered. This suggests that the variability of the topology of gap-junction couplings could play a role in the variability of the ICW propagation range. Since this hypothesis is very difficult to investigate with current experimental approaches, we explored it using a biophysically realistic model of three-dimensional astrocyte networks in which we varied the topology of the astrocyte network, while keeping intracellular properties and spatial cell distribution and density constant. Computer simulations of the model suggest that changing the topology of the network is indeed sufficient to reproduce the distinct ranges of ICW propagation reported experimentally. Unexpectedly, our simulations also predict that sparse connectivity and restriction of gap-junction couplings to short distances should favor propagation while long–distance or dense connectivity should impair it. Altogether, those results provide support to recent experimental findings that point towards a significant functional role of the organization of gap-junction couplings into proper astroglial networks. Dynamic control of this topology by neurons and signaling molecules could thus constitute a new type of regulation of neuron-glia and glia-glia interactions.

This result has been published in [18] and as conference talks. It is based on J. Lallouette's PhD thesis work in collaboration with M. De Pittà (postdoc in the team) and E Ben-Jacob, Tel Aviv University, Israel.

5.3. Glutamate Mediated Astrocytic Filtering of Neuronal Activity

Participants: H. Berry, J. Lallouette, M. De Pittá

Neuron-astrocyte communication is an important regulatory mechanism in various brain functions but its complexity and role are yet to be fully understood. In particular, the temporal pattern of astrocyte response to neuronal firing has not been fully characterized. Here, we used neuron-astrocyte cultures on multi-electrode arrays coupled to Ca2+ imaging and explored the range of neuronal stimulation frequencies while keeping constant the amount of stimulation. Our results reveal that astrocytes specifically respond to the frequency of neuronal stimulation by intracellular Ca2+ transients, with a clear onset of astrocytic activation at neuron firing rates around 3-5 Hz. The cell-to-cell heterogeneity of the astrocyte Ca2+ response was however large and increasing with stimulation frequency. Astrocytic activation by neurons was abolished with antagonists of type I metabotropic glutamate receptor, validating the glutamate-dependence of this neuron-to-astrocyte pathway. Using a realistic biophysical model of glutamate-based intracellular calcium signaling in astrocytes, we suggest that the stepwise response is due to the supralinear dynamics of intracellular IP3 and that the heterogeneity of the responses may be due to the heterogeneity of the astrocyte-to-astrocyte couplings via gap junction channels. Therefore our results present astrocyte intracellular Ca2+ activity as a nonlinear integrator of glutamate-dependent neuronal activity.

This result has been published in a paper currently in press, [26] and is a direct result from J. Lallouette's PhD thesis in collaboration with Y. Hanein's group, in Tel Aviv University (for the experimental measurements), M. De Pittà (postdoc in the team), and E Ben-Jacob, Tel Aviv University, Israel.

5.4. Space-induced bifurcation in repression-based transcriptional circuits

Participants: H. Berry, A. Lo Van

Experimental measurements of the mobility of macromolecules, especially proteins, in cells and their membranes consistently report transient subdiffusion with possibly position-dependent —non-homogeneous— properties. However, the spatiotemporal dynamics of protein mobility when transient subdiffusion is restricted to a subregion of space is still unclear. We have investigated the spatial distribution at equilibrium of proteins undergoing transient subdiffusion due to continuous-time random walks (CTRW) in a restricted subregion of a two-dimensional space. Our Monte-Carlo simulations suggest that this process leads to a non-homogeneous spatial distribution of the proteins at equilibrium, where proteins increasingly accumulate in the CTRW subregion as its anomalous properties are increasingly marked. These results suggest that, even though they exhibit the same time-dependence of the mean-squared displacement, the different scenarios proposed to account for subdiffusion in the cell lead to different protein distribution in space, even at equilibrium and without coupling with reaction. We also we assessed the influence of the spatial distribution of the genes on the dynamics of 3-gene transcriptional ring networks regulated by repression, i.e. repressilator circuits. Our simulations suggest that variations of spatial parameters – namely the degree of demixing of the positions of the gene or the spatial range of the mRNA and proteins (i.e. the typical distance they travel before degradation) – have dramatic effects by switching the dynamical regime from spontaneous oscillations to a stationary state where each species fluctuates around a constant value. By analogy with the bifurcations arising from the variation of kinetic parameters, we referred to those transitions as space-induced bifurcations. Therefore, our results strongly support the idea that the spatial organization of the molecular actors of transcriptional networks is crucial for the dynamics of gene expression and suggest that the spatial localization of the synthetic genes in the cell could be used as an additional toggle to control the dynamics of the inserted construct in synthetic biology experiments.

This group of results has been published in [20], [13], [12] and [23]. It consists in the PhD and Master works of B. Caré and A. Lo Van, respectively, and a collaboration with H Chaté, CEA, Saclay.

5.5. Modeling interaction of transcription processes in neighbour genes

Participants: G. Beslon, S. Meyer

During the transcription process, the genetic sequence encoded in the DNA molecule is expressed by an enzymatic complex. This process is often considered as independent for each gene, despite numerous reported cases of one transcribed gene perturbing a neighbour gene's expression, which is then regarded as a side-effect. Here, we suggest in the contrary that such interactions are a widespread feature, resulting from the propagation along the DNA molecule of mechanical stress generated during gene transcription. This torsional stress modifies the facility with which the transcription machinery separates the two strands of the double-helix in order to access the bases, and thus the expression level of any gene located nearby. We develop a quantitative model of this effect, showing that it depends strongly on the orientation of the genes, which is confirmed by the analysis of in vivo expression levels in the drosophila genome. This observation suggests that torsional coupling may play an important role in genetic regulation, and might favor the orientation-dependent co-localization of genes involved in similar functions, which need to be expressed together.

Publication: [21]

5.6. A model of genome size evolution

Participants: G. Beslon, C. Knibbe, S. Fisher

Even though numerous genomic sequences are now available, evolutionary mechanisms that determine genome size, notably their fraction of non-coding DNA, are still debated. In particular, although several mechanisms responsible for genome growth (proliferation of transposable elements, gene duplication and divergence, etc.) were clearly identified, mechanisms limiting the overall genome size remain unclear.

In collaboration with Samuel Bernard (Inria Dracula Team and Institut Camille Jordan, UMR CNRS 5208, Lyon), we have developed a model for genome size evolution that takes into account both local mutations such as small insertions and small deletions, and large chromosomal rearrangements such as duplications and large deletions. We introduced the possibility of undergoing several mutations within one generation. The model, albeit minimalist, revealed a non-trivial spontaneous dynamics of genome size: in the absence of selection, an arbitrary large part of genomes remains beneath a finite size, even for a duplication rate 2.6-fold higher than the rate of large deletions, and even if there is also a systematic bias toward small insertions compared to small deletions. Specifically, we showed that the condition of existence of an asymptotic stationary distribution for genome size non-trivially depends on the rates and mean sizes of the different mutation types. We also gave upper bounds for the median and other quantiles of the genome size distribution, and argue that these bounds cannot be overcome by selection. Taken together, these results show that the spontaneous dynamics of genome size naturally prevents it from growing infinitely, even in cases where intuition would suggest an infinite growth. This work was part of Stephan Fischer's PhD thesis, which was defended in December 2013.

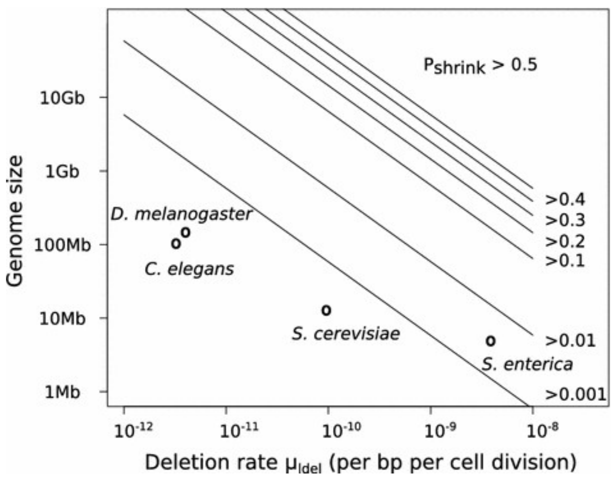


Figure 2. Comparison of the bounds on genome size with the genome size for four organisms. Spontaneous deletion rates were computed per base pair and per cell division from experimental data on mutation accumulations for the bacterium Salmonella enterica, the budding yeast Saccharomyces cerevisiae, the worm Caenorhabditis elegans and the fruit fly Drosophila melanogaster. The value next to each line is the lower bound for the probability that a genome located along this line will shrink at the next step in our model for equal duplication and deletion rates.

This year, using quantitative numerical examples with parameters taken from biological data, we showed that, in practice, a shrinkage bias appears very quickly in genomes undergoing mutation accumulation, even though DNA gains and losses appear to be perfectly symmetrical at first sight. This spontaneous dynamics provides the genome with a stability-related size limit below which it can be influenced by other evolutionary forces (selection, drift, biases, ...).

All this work has been published this year [15], and is already mentioned as "most read article" by Springer.

5.7. A novel view on reductive evolution

Participants: G. Beslon, C. Knibbe, B. Batut

Bacterial genomes show substantial variations in size. The smallest bacterial genomes are those of endocellular symbionts of eukaryotic hosts, which have undergone massive genome reduction and show patterns that are consistent with the degenerative processes that are predicted to occur in species with small effective population sizes. However, similar genome reduction is found in some free-living marine cyanobacteria that are characterized by extremely large populations. Using a combination of bioinformatics approaches and of in silico experimental evolution (with the aevol model), we have been able to propose a scenario that explains the reductive evolution of marine bacteria.

This work was part of Bérénice Batut's PhD thesis [10], which was defended in November 2014. Bérénice was co-supervised by Guillaume Beslon and Carole Knibbe (Inria BEAGLE team) for the simulations and by Gabriel Marais and Vincent Daubin (Laboratoire de Biométrie et Biologie Evolutive, UMR CNRS 5558) for the genomic analyses. This work had already yielded a publication in 2013 [34]. This year, we published a review in the high-level journal [11]. The scenario proposed in the PhD manuscript, as well as the simulations and analyses done this year to support it, should be published in 2015.

5.8. Genome evolution aware gene trees

Participant: E. Tannier

Traditionally the inference of a gene tree is made from a multiple alignment of homologous sequences according to a model of molecular evolution. Trees for several gene families are thus constructed one by one, independently from each other. Constructed this way trees often carry unresolutions or bad resolutions. Information for their full resolution may lie in the poorly exploited dependency between gene families, each bringing information for the resolution of the others. We used several kinds of such dependencies in the construction of gene trees: information from a species tree through a model of gene content evolution, information from extant synteny through ortholog predictions, and information from ancestral synteny through a model of gene neighborhood evolution. We developed, improved, implemented and gave a user interface to several "correction" techniques, yielding a series of correction modules called "RefineTree". We tested its parts on simulated data and apply it on the full set of gene families from the Ensembl Compara database. We showed that according to several measures including the tree likelihood computed from sequence evolution, the stability of genome content and the linearity of ancestral chromosomes, trees corrected by refineTree are arguably more plausible than the ones stored by Ensembl.

This work has been achieved by Magali Semeria, Laurent Gueguen (LBBE) and Eric Tannier in Lyon, in collaboration with Nadia El-Mabrouk's group from the computer science department of the university of Montreal. This collaboration started when Nadia El-Mabrouk was an Inria visiting professor in our team in 2012 and 2013. An article has been submitted.

5.9. Variable food availability increases weight: a mathematical prediction

Participant: H. Soula

Due to the conservation of energy, the energy storage in adipose tissue reflect the difference of energy expenditure and energy intake. Without change in physical activity, the main paradigm has always been that this storage does not depend on the timing of intake but on its whole temporal integration: the overall food intake. However, mammal and especially rats can compensate energy expenditure to save energy in case of starving. This adaptation should provoke variation in energy expenditure when food availability varies in time. Using animal experiments and mathematical modelling, we showed that indeed food availability variation - while conserving the same amount of energy - can disrupt and perturb energy balance. Submitted to variation in availability with a period above 4 weeks, rats where bigger with higher fat mass than control. Even so these rats had eaten the same amount of food as the control group during the same period. Our mathematical model uses delay equations and can predict both the food intake and the body weight variations. We showed that delay in energy saving adaptation cause this variation and estimate the lag at 1 week. This result could very well apply to humans in the so called 'yoyo regime'. Regime that are stopped are a typical case of food intake variation and could cause greater fat accretion instead of body weight reduction. We show that this should happen if the regime lasts longer than one week.

This result has been the subject of an article in the weekly journal of Inserm Rhônes-Alpes with an interview of author H. Soula.

5.10. Insights on gene family dynamics from digital genetics experiments

Participants: C. Knibbe

Gene families are sets of homologous genes formed by duplications of a single original gene. Inferring their history in terms of gene duplications, gene losses and gene mutations yields fundamental insights into the molecular basis of evolution. However, the traditional approach, the phylogenetic inference of gene family evolution, faces two difficulties: (i) the delimitation of gene families based on sequence similarity, and (ii) the fact that the models of evolution used for reconstruction are tested against simulated data that are produced by the model itself. This year, we showed that digital genetics, or in silico experimental evolution, can provide thought-provoking synthetic gene family data, robust to rearrangements in gene sequences and, most importantly, not biased by where and how we think natural selection should act. Using Aevol, we analyzed the evolution of 3,512 synthetic gene families under directional selection. The turnover of gene families in evolutionary runs was such that only 21% of those families would be accessible for classical phylogenetic inference. Extinct families showed patterns different from the final, observable ones, both in terms of dynamics of gene gains and losses and in terms of gene sequence evolution. This study also reveals that gene sequence evolution, and thus evolutionary innovation, occurred not only through local mutations, but also through chromosomal rearrangements that re-assembled parts of existing genes.

This work was published in the international conference ALIFE 2014 [28].

BIGS Project-Team

5. New Results

5.1. Analysis of high dimensional data

Participants: K. Duarte, S. Ferrigno, J.-M. Monnez, A. Muller-Gueudin, S. Tindel

5.1.1. Online partial principal component analysis of a data stream

Consider a data stream and suppose that each data vector is a realization of a random vector whose expectation varies with time, the law of the centered data vector being stationary. Consider the principal component analysis (PCA) of this centered vector called partial PCA. In this study are defined online estimations of the first principal axes by stochastic approximation processes using a data batch at each step of the process or all the data until the current step. This extends a former result obtained by J.-M. Monnez by using one data vector at each step. This is applied to partial generalized canonical correlation analysis by defining a stochastic approximation process of the metric involved in this case using all the data until the current step. If the expectation of the data vector varies according to a linear model, a stochastic approximation process of the model parameters is used. All these processes can be performed in parallel. A forthcoming preprint by R. Bar and J.-M. Monnez will discuss those aspects.

5.1.2. Data analysis for cumulative exposure Index

Everyone is subject to environmental exposures from various sources, with negative health impacts (air, water and soil contamination, noise ...) or with positive effects (e.g., green space). Studies considering such complex environmental settings in a global manner are rare. In [5] we propose to use statistical factor and cluster analyses to create a composite exposure index with a data-driven approach, in view to assess the environmental burden experienced by populations. The study was carried out in the Great Lyon area (France, 1.2M inhabitants) at the census block group (BG) scale. We used as environmental indicators ambient air NO2 annual concentrations, noise levels, proximity to green spaces, to industrial plants, to polluted sites and to road traffic. Although it cannot be applied directly for risk or health effect assessment, the resulting index can help to identify hot spots of cumulative exposure, to prioritize urban policies or to compare the environmental burden across study areas in an epidemiological framework.

5.1.3. A simultaneous stepwise covariate selection

In supervised learning the number of values of a response variable to predict can be high. Also clustering them in a few clusters can be useful to perform relevant supervised classification analysis. On the other hand selecting relevant covariates is a crucial step to build robust and efficient prediction models, especially when too many covariates are available in regard to the overall sample size. As a first attempt to solve these problems, we had already devised in a previous study an algorithm that simultaneously clusters the levels of a categorical response variable in a limited number of clusters and selects forward the best covariates by alternate minimization of Wilks's Lambda. In the project carried out this year, we first extend the former version of the algorithm to a more general framework where Wilks's Lambda can be replaced by any model selection criterion. We also turned forward selection into stepwise selection in order to remove covariates in real time if necessary. Finally an application of our algorithm to real datasets from peanut allergy studies allowed to get confirmation of some previously published results and suggested new discoveries. The possibilities of this algorithm are promising and it is hoped to be useful for many practitioners.

5.1.4. Prognostic value of the Strauss estimated plasma

We describe here an application oriented study lead jointly by J.-M. Monnez and a medical team under the supervision of E. Albuisson at CHU Brabois. The objective is to assess the prognostic value of estimations of volemia, or of their variations, beyond clinical examination in a post-hoc analysis of the Eplerenone Post-Acute Myocardial Infarction (AMI) Heart Failure (HF) Efficacy and Survival Study (EPHESUS). Assessing congestion post-discharge is indeed challenging but of paramount importance to optimize patient management and prevent hospital readmissions. The analysis was performed in a subset on 4957 patients with available data (within a full dataset of 6632 patients). Study endpoint was cardiovascular death and/or hospitalization for HF between month 1 and month 3 after post-AMI HF. Estimated plasma volume variations between baseline and month 1 were estimated by the Strauss formula, which includes hemoglobin and hematocrit ratios. Other potential predictors including congestion surrogates, hemodynamic and renal variables, and medical history variables were tested. An instantaneous estimation of plasma volume at month 1, ePVS M1, was defined and also tested. Multivariate analysis was performed using stepwise logistic regression and linear discriminant analysis. In HF complicating MI, congestion assessed by the Strauss formula and an instantaneous derived measurement of plasma volume displayed an added predictive value of early cardiovascular events, beyond routine clinical assessment. Trials assessing congestion management guided by this simple tool to monitor plasma volume are warranted.

5.1.5. Non parametric estimation of the conditional cumulative distribution function

This project fits into the global aim of improving local regression techniques. Indeed, we propose in [21] to study the local linear estimator of the conditional distribution function. Namely, having an i.i.d. sample $(X_i, Y_i)_{1\leq i\leq n}$, we estimate the conditional distribution function $F(t|x) = \mathbb{P}(Y \leq t|X = x)$ by:

$$\widehat{F}_n^{(1)}(t, h_n|x) = \frac{\widehat{f}_{n,2}(x, h_n)\widehat{r}_{n,0}(x, t, h_n) - \widehat{f}_{n,1}(x, h_n)\widehat{r}_{n,1}(x, t, h_n)}{\widehat{f}_{n,0}(x, h_n)\widehat{f}_{n,2}(x, h_n) - \left(\widehat{f}_{n,1}(x, h_n)\right)^2} \tag{1}$$

where $^{(1)}$ denotes the order 1 of the local polynomial estimator, $\widehat{f}_{n,j}$ stands for a kernel estimator with order j of the probability density function f_X of X, $\widehat{r}_{n,j}$ estimates the distribution of the couple (X, Y) and h_n is a bandwidth parameter.

This estimator is a particular case of the local polynomial estimators. It is the local polynomial estimator of order $p = 1$. Another simpler estimator, with order $p = 0$, is well known as the Nadaraya-Watson estimator.

We are interested in showing the advantage of this estimator over the Nadaraya-Watson estimator. We show asymptotic results for our estimator (exact rate of uniform consistency), and establish also uniform asymptotic certainty bands for the conditional cumulative distribution function.

We obtain the following result under some assumptions on the cumulative distribution F, f_X, the kernel K and the bandwidth h_n,

$$\sup_{t\in\mathbb{R}}\sup_{x\in I}\sqrt{\frac{nh_n}{\log(h_n^{-1})}}\left|\widehat{F}_n^{(1)}(t, h_n|x) - \widehat{\mathbb{E}}\left(\widehat{F}_n^{(1)}(t, h_n|x)\right)\right| \underset{n\to+\infty}{\overset{\mathbb{P}}{\rightarrow}} \sigma_F(I) \tag{2}$$

where

$$\sigma_F^2(I) = \frac{||K||_2^2}{2\inf_{x\in I} f_X(x)}.$$

As corollaries of this result, we extend our results to other statistical functions, such as the quantiles and the regression function.

We illustrate our results with simulations and an application on foetopathologic data.

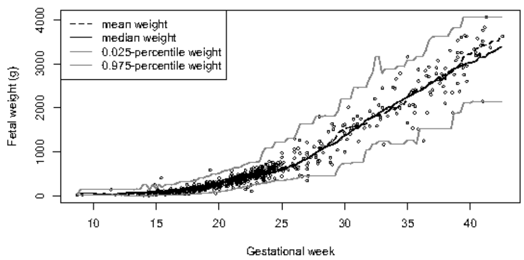


Figure 5. Fetal weight during the pregnancy: estimation of mean and quantiles from our local polynomial regression method.

We have also started a study about the regression function in the application on foetopathologic data. We consider the nonparametric model

$$Y = m(X) + \epsilon,$$

where Y is the fetal weight , X are the gestational weeks, m is a smooth unknown function and ϵ the error. The goal is to provide a test to detect significant features (or change points) of this regression curve. The regression curve is estimated using local polynomial kernel smoothers.

5.2. Stochastic modeling for complex and biological systems

Participants: R. Azaïs, T. Bastogne, C. Lacaux, A. Muller-Gueudin, S. Tindel, P. Vallois, S. Wantz-Mézières

5.2.1. Modelisation of networks of multiagent systems

We relate here the beginning of collaboration between A. Gueudin, R. Azaïs and some automatic control researchers in Nancy.

We consider networks, modeled as a graph with nodes and edges representing the agents and their interconnections, respectively. The connectivity of the network, persistence of links and interactions reciprocity influence the convergence speed towards a consensus.

The problem of consensus or synchronization is motivated by different applications as communication networks, power and transport grids, decentralized computing networks, and social or biological networks.

We then consider networks of interconnected dynamical systems, called agents, that are partitioned into several clusters. Most of the agents can only update their state in a continuous way using only inner-cluster agent states. On top of this, few agents also have the peculiarity to rarely update their states in a discrete way by resetting it using states from agents outside their clusters. In social networks, the opinion of each individual evolves by taking into account the opinions of the members belonging to its community. Nevertheless, one or several individuals can change its opinion by interacting with individuals outside its community. These inter-cluster interactions can be seen as resets of the opinions. This leads us to a network dynamics that is expressed in term of reset systems. We suppose that the reset instants arrive stochastically following a Poisson renewal process.

5.2.2. Tumor growth modeling

A cancer tumor can be represented for simplicity as an aggregate of cancer cells, each cell behaving according to the same discrete model and independently of the others. Therefore to measure its size evolution, it seems natural to use tools coming from dynamics of population, for instance the logistic model. This deterministic framework is well-known but the stochastic one is worthy of interest. We are currently studying in [22] a model in which we suppose that the size V_t at time t of the tumor is a diffusion process of the type :

$$\begin{cases} dV_t = r\,V_t\,\left(1 - \frac{V_t}{\kappa}\right) - c\,V_t + \beta\,V_t\,dB_t \\ V_0 = v > 0 \end{cases} \tag{3}$$

where $(B_t)_{t \geq 0}$ is a standard brownian motion starting from zero.Then (i) We define a family of time continuous Markov chains which models the evolution of the rate of malignant cells and approximate (under some conditions) the diffusion process (V_t). (ii) We study in depth the solution to equation (3). This diffusion process lives in a domain delimited by two boundaries: 0 and $\kappa > 0$. In this stochastic setting, the role of κ is not so clear and we contribute to understand it. We describe the asymptotic behavior of the diffusion according to the values of the parameters. The tools we resort to are boundary classification criteria and Laplace transform of the hitting time to biological worthwhile level. We are able in particular to express the mean of the hitting time.

5.2.3. Anisotropic random fields

Hermine Biermé (Tours) and Céline Lacaux follow in [19] their collaboration in the study of anisotropic random fields. They have extended their previous works in the framework of conditionally sub-Gaussian random series. For such anisotropic fields, they have obtained a modulus of continuity and a rate of uniform convergence. Their framework allows to study e.g., Gaussian fields, stable random fields and multi-stable random fields.

5.2.4. Inference for dynamical systems driven by Gaussian noises.

As mentioned in the *Scientific Foundations* Section, the problem of estimating the coefficients of a general differential equation driven by a Gaussian process is still largely unsolved. To be more specific, the most general ($\mathbb{R}$-valued) equation handled up to now as far as parameter estimation is concerned is of the form:

$$X_t^\theta = a + \theta \int_0^t b(X_u)\,du + B_t,$$

where θ is the unknown parameter, b is a smooth enough coefficient and B is a one-dimensional fractional Brownian motion. In contrast with this simple situation, our applications of interest (see the *Application Domains* Section) require the analysis of the following $\mathbb{R}^n$-valued equation:

$$X_t^\theta = a + \int_0^t b(\theta; X_u)\,du + \int_0^t \sigma(\theta; X_u)\,dB_t, \tag{4}$$

where θ enters non linearly in the coefficient, where σ is a non-trivial diffusion term and B is a d-dimensional fractional Brownian motion. We have thus decided to tackle this important scientific challenge first.

To this aim, here are the steps we have focused on in 2014:

- A better understanding of the underlying rough path structure for equation (4). This includes two studies on differential systems driven by some general Gaussian noises in infinite dimensions: [17] on the Parabolic Anderson model, and [16] about viscosity solutions in the rough paths setting.
- Study of densities for general systems driven by Gaussian noises as in [18] and [15].
- Ergodic aspects, which are another important ingredient for estimation procedures for stochastic differential equations, are handled in [3].

5.2.5. Extremal process

In extreme value theory, one of the major topics is the study of the limiting behavior of the partial maxima of a stationary sequence. When this sequence is i.i.d., the unique limiting process is well-known and called the extremal process. Considering a long memory stable sequence, the limiting process is obtained as a simple power time change extremal process. Céline Lacaux and Gennady Samorodnistky have proved in [23] that this limiting process can also be interpreted as a restriction of a self-affine random sup measure. In addition, they have established that this random measure arises as a limit of the partial maxima of the same long memory stable sequence, but in a different space. Their results open the way to propose new self-similar processes with stationary max-increments.

5.2.6. Self-nested structure of plants

In a recent work, Godin and Ferraro designed a method to compress tree structures and to quantify their degree of self-nestedness. This method is based on the detection of isomorphic subtrees in a given tree and on the construction of a DAG, equivalent to the original tree, where a given subtree class is represented only once (compression is based on the suppression of structural redundancies in the original tree). In the compressed graph, every node representing a particular subtree in the original tree has exactly the same height as its corresponding node in the original tree.

The degree of self-nestedness is defined as the edit-distance between the considered tree structure and its nearest embedded self-nested version. Indeed, finding the nearest self-nested tree of a structure without more assumptions is conjectured to be an NP-complete or NP-hard problem. We thus design a heuristic method based on interacting simulated annealing algorithms to tackle this difficult question. This procedure is also a keystone in a new topological clustering algorithm for trees that we propose in this work. In addition, we obtain new theoretical results on the combinatorics of self-nested structures. For instance, we have shown that the number $C_{\leq H}(m)$ of self-nested trees with maximal height H and a ramification number for each vertex less than m satisfies the following formula,

$$C_{\leq H}(m) = \sum_{h=1}^{H} \prod_{i=1}^{h} \binom{m+h-i}{h-i+1}.$$

In particular, the cardinality $C_{=h}(m)$ of self-nested trees with exact height h evolves according to

$$\log C_{=h}(m) \sim \frac{(m+h)^2}{2} \log(m+h) - \frac{h^2}{2} \log h - \frac{m^2}{2} \log m - mh \log m,$$

when m and h simultaneously go to infinity. The redaction of an article is currently in progress.

5.2.7. Semi-parametric inference for a growth-fragmentation model

Statistical inference for piecewise-deterministic Markov processes has been extensively investigated for a few years under some ergodicity conditions. Our paper [2] is dedicated to a statistical approach for a particular non ergodic growth-fragmentation model for which the set $[0,1]$ is absorbing. This kind of stochastic process may model the dynamic of a malthusian population for which there exists an extinction threshold. We focus on the estimation of the extinction probability and of the distribution of the extinction time from only one path of the model within a long time interval.

We establish that the absorption probability p is the unique solution in an appropriate space of a Fredholm equation of the second kind whose parameters are unknown,

$$p - Kp = s,$$

where K is an integral operator depending explicitly on the main features of the model. From n data, we estimate this important characteristic of the underlying process by solving numerically the estimated Fredholm equation. Indeed, $\widehat{p}_{n,m}$ is defined as the approximated solution of the equation $p - \widehat{K}_n p = \widehat{s}_n$ after m steps of the algorithm. Fortunately, this procedure allows us to estimate also the extinction time.

We have shown the convergence in probability of the proposed estimators under some usual asymptotic conditions. In particular, we have,

$$\forall \varepsilon > 0, \; \mathbf{P}\left(\|p - \widehat{p}_{n,m}\|_1 > \varepsilon\right) \to 0,$$

when n and m simultaneously go to infinity. The good behavior of our estimates on finite sample sizes is presented in Figure 6 . In future works, we plan to apply this methodology to more intricate situations, in particular for the pharmacokinetics and pharmacodynamics stochastic models recently introduced in the literature.

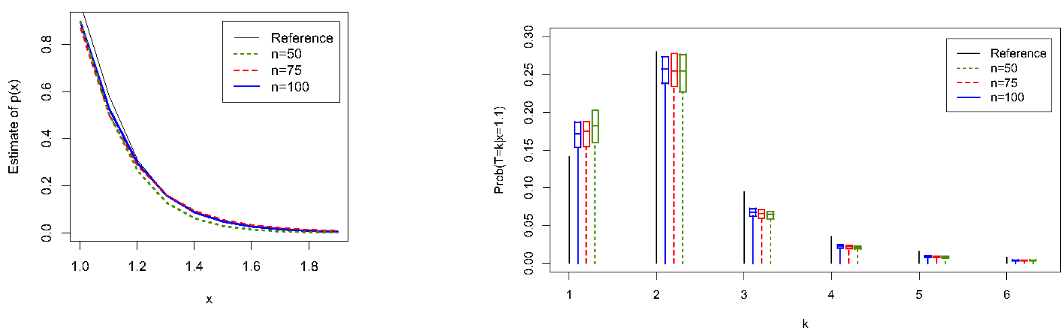


Figure 6. Probability of extinction $p(x)$ and its estimates for a trajectory starting from the initial population $1 \leq x \leq 2$ (left) and distribution of the extinction time and its estimates for a trajectory starting from $x = 1.1$ (right).

5.2.8. A Model-based Pharmacokinetics Characterization Method of Engineered Nanoparticles for Pilot Studies

Recent developments on engineered multifunctional nanomaterials have opened new perspectives in oncology. But assessment of both quality and safety in nanomedicine requires new methods for their biological characterization. We have recently proposed a new model-based approach for the pre-characterization of multifunctional nanomaterials pharmacokinetics in small scale in vivo studies. Two multifunctional nanoparticles, with and without active targeting, designed for photodynamic therapy guided by magnetic resonance imaging are used to exemplify the presented method. It allows the experimenter to rapidly test and select the most relevant pharmacokinetic (PK in the sequel) model structure planned to be used in the subsequent explanatory studies. We also show that the model parameters estimated from the in vivo responses provide relevant preliminary information about the tumor uptake, the elimination rate and the residual storage. For some parameters, the accuracy of the estimates is good enough to compare and draw significant pre-conclusions. A third advantage of this approach is the possibility to optimally refine the in vivo protocol for the subsequent explanatory and confirmatory studies complying with the 3Rs (reduction, refinement, replacement) ethical recommendations. More precisely, we show that the identified model may be used to select the appropriate duration of the magnetic resonance imaging sessions planned for the subsequent studies. The proposed methodology integrates magnetic resonance image processing, continuous-time system identification algorithms and statistical analysis. Except, the choice of the model parameters to be compared and interpreted, most of the processing procedure may be automated to speed up the PK characterization process at an early stage of experimentation.

More specifically, our efforts have been split into the following tasks:

- The article [6] gives an application of statistical methods for the design of experiments to optimize the formulation of a composite molecule in photodynamic therapy. The associated know-how has been transferred to the start-up CYBERnano to be generalized to the rational design of engineered nanoparticles. Collaboration with CRAN and LRGP (Nancy) and UNINE (Neuchâtel, Suisse).
- In [12], in vivo application of photodynamic therapy, a mathematical model and computational simulations of the light propagation in biological tissues were developed to help biologists to determine *a priori* some parameters of the experimental protocol. More precisely, the numerical results were used to select the most suited position of the optical fiber to be implemented within the animal brain. This equipment is required to bring the light and thus activate the molecule within the tumor. The therapeutical objective was to maximize the homogeneity of light intensity within the tumor volume.
- Obstacles and challenges to the clinical use of the photodynamic therapy (PDT) are numerous: large inter-individual variability, heterogeneity of therapeutic predictability, lack of in vivo monitoring concerning the reactive oxygen species (ROS) production, etc. All of these factors affect in their ways the therapeutic response of the treatment and can lead to a wild uncertainty on its efficiency. To deal with these variability sources, we have designed and developed an innovative technology able to adapt in realtime the width of light impulses during the photodynamic therapy. The first objective is to accurately control the photobleaching trajectory of the photosensitizer during the treatment with a subsequent goal to improve the efficacy and reproducibility of this therapy. In this approach, the physician a priori defines the expected trajectory to be tracked by the photosensitizer photobleaching during the treatment. The photobleaching state of the PS is regularly measured during the treatment session and is used to change in real-time the illumination signal. This adaptive scheme of the photodynamic therapy has been implemented, tested and validated during in vitro tests. These tests show that controlling the photobleaching trajectory is possible, confirming the technical feasibility of such an approach to deal with inter-individual variabilities in PDT. These results, contained in [13], open new perspectives since the illumination signal can be different from a patient to another according to his individual response. This study has proven its interest by showing promising results in an in vitro context, which has to be confirmed by the current in vivo experiments. However, it is fair to say that in a near future, the proposed solution could lead, in fine, to an optimized and personalized PDT. A patent was deposited subsequently. Collaboration with CRAN (Nancy).
- The communications [8], [9] and [10] present successful applications of a model-based design of nanoparticles. This approach is based on statistical design of experiments and black-box modeling in cell biology. The associated know-how has been transferred to the start-up CYBERnano. Collaboration with CEA LETI and INSERM (Grenoble).

BIOCORE Project-Team

6. New Results

6.1. Highlights of the Year

- We reanalyzed the so-called Marginal Value Theorem (MVT), first published in 1976, in a paper published in Ecology Letters [23]. This theorem, also used in human behavior and economics, establishes how individuals should behave to optimize resource exploitation. Despite the thousands of papers written on the subject, we obtained the first mathematical characterization of how habitat characteristics affect the optimal foraging strategy. Mathematical foundations for this work were given in [24].
- The analysis of metabolic networks is generally made under the assumption (so called "balanced growth") that there is no internal accumulation of metabolites. However, this hypothesis is clearly wrong for microalgae, which store lipids and carbohydrates during the day and consume it during the night. A new formalism, called DRUM (Dynamic Reduction of Unbalanced Metabolism) was developed [16], assuming that the balanced growth is valid only in subnetworks, but that there can be accumulation between these modules (which often represent spatial distribution in the cell). This approach was successfully used to represent the dynamics of carbon accumulation in the microalgae *Tisochrysis lutea* under light/dark cycles, or in response to a nitrogen starvation. It also well described the diauxic heterotrophic growth of *Chlorella pyrernoidosa* [11].

6.2. Mathematical methods and methodological approach to biology

6.2.1. Mathematical analysis of biological models

6.2.1.1. Mathematical study of semi-discrete models

Participants: Jean-Luc Gouzé, Frédéric Grognard, Ludovic Mailleret, Pierre Bernhard, Elsa Rousseau, Nicolas Bajeux, Bapan Ghosh.

Semi-discrete models have shown their relevance in the modeling of biological phenomena whose nature presents abrupt changes over the course of their evolution [99]. We used such models and analyzed their properties in several practical situations that are developed in Section 6.3.2 , some of them requiring such a modeling to describe external perturbations of natural systems, and others to take seasonality into account. External perturbations of interacting populations occur when some individuals are introduced or removed from a natural system, which occurs frequently in pest control applications, either through the direct removal of pests [62], or through the introduction of biological control agents [45], [60], [54]. Seasonality is an important property of most agricultural systems in temperate environments since the year is divided into a cropping season and a ‘winter’ season, where the crop is absent, as in our analysis of eco-evolutionary dynamics of plant pathogens [25], [59]

6.2.1.2. Model reduction and sensitivity analysis

Participant: Suzanne Touzeau.

Dynamic models representing complex biological systems with numerous interactions can reach high dimensions and include complex nonlinearities. Especially if data are scarce, identifying the model parameters is then a challenge. So we designed an ad-hoc method based on global sensitivity analysis to simplify the model and determine the most influential parameters. It was applied to a within-host immunological model [30], [61]. This application was part of Natacha Go’s PhD thesis, supervised by S. Touzeau and C. Belloc (BioepAR, INRA & Oniris Nantes) [90].

6.2.2. Metabolic and genomic models

Participants: Jean-Luc Gouzé, Madalena Chaves, Alfonso Carta, Ismail Belgacem, Olivier Bernard, Caroline Baroukh, Jean-Philippe Steyer, Diego de Pereda Sebastian, Francis Mairet.

6.2.2.1. Continuous models analysis

Transcription and translation models in bacteria We study detailed models of transcription and translation for genes in a bacterium, in particular the model of gene expression of RNA polymerase. With techniques of monotone systems, and time scale hypotheses, we can show the stability of the fast part of these systems, and reduce them to much smaller models [49], [48], [47]. We also study other models of the global cellular machinery. This is part of the PhD theses of Ismael Belgacem and Alfonso Carta [12], and done in collaboration with Inria IBIS project-team.

A model of synthesis of a virulence factor In collaboration with J.-A. Sepulchre (INLN Nice), we model the production of a virulence factor by a bacterium in a continuous stirred tank reactor. The production of this enzyme is genetically regulated, and degrades a polymeric external substrate into monomers.

Analysis and reduction of biochemical models In collaboration with D. Ropers (Inria IBIS project team), we address the problem of reduction of large biochemical networks, to decompose the dynamic behavior of the whole system into simpler models. This is the subject of the thesis of S. Casagranda.

Design of a bistable switch to control cellular uptake In joint work with Diego Oyarzún (Imperial College), we explore the idea of constructing a synthetic bistable system using an unbranched metabolic chain with a global enzyme regulator. Bistability can be achieved by choosing an appropriate pattern of regulation and deriving conditions on the promoter dynamic ranges to guarantee a bistable uptake flux. This work started during the visit of Diego to Biocore in October 2014.

Analysis of signaling pathways leading to apoptosis In joint work with Jérémie Roux (Marie Curie Fellow, IRCAN Nice), a cascade of signaling modules leading to apoptosis (or programmed cell death) was implemented and studied through simulations. The goal of this work is to determine whether, and at which stage in the pathway, the system may exhibit bistability. This was the work of Xiao Han's internship.

6.2.2.2. Hybrid models analysis

Piecewise quadratic systems for studying growth rate in bacteria The class of piecewise affine systems was extended to deal with dynamics dependent on dilution due to cell growth rate, leading to switched-piecewise quadratic systems [85]. These new systems use an expression for growth rate that may depend on any number of variables and have several quadratic modes. The behavior of piecewise quadratic systems introduces new features, notably regarding solutions at the thresholds when the vector fields are opposing: not only sliding mode solutions but also oscillatory behavior may happen. Part of this work is in the PhD thesis of Alfonso Carta [12].

Attractor computation using interconnected Boolean networks The method developed in [10] has been extended towards a better characterization of the attractors of the interconnected system in terms of invariant sets [26]. The method was used to test growth rate models in E. Coli using Boolean networks.

Analysis of circadian rhythms in cyanobacteria The model describing the system responsible for the circadian rhythm of cyanobacteria previously proposed in [86] has been improved in [50]. Here, we have tested the robustness of the circadian rhythm with respect to the perturbations inherent to the noisy environment of the cell, including cell growth and division. The interconnection between two models was studied: circadian rhythm and a stochastic model for cell division.

Structure estimation for Boolean models of gene regulation networks The problem of estimating Boolean models of gene networks from few and noisy measurements is addressed in [84], joint work with C. Breindl and F. Allgöwer from the University of Stuttgart. The class of unate or canalizing Boolean functions has been further considered and represented by multi-affine polynomials, leading to a reformulation of the estimation problem as a mixed integer linear program.

Structural principles for the existence of limit cycles in two-dimensional piecewise affine models Using concavity and continuity properties of Poincaré maps, we have derived some structural principles which link the topology of the transition graph to the existence, number and stability of limit cycles in a class of two-dimensional piecewise affine biological models of genetic networks [14].

6.2.2.3. Estimation and control

Optimal allocation of resources in a bacterium We study by techniques of optimal control the optimal allocation between metabolism and gene expression during growth of bacteria [52], in collaboration with Inria IBIS project-team.

Estimation of biological models In a joint work with Diego de Pereda (visiting PhD student), we studied observers and interval observers for models of glucose concentration in diabetes.

6.3. Fields of application

6.3.1. Bioenergy

6.3.1.1. Modelling of microalgae production

Participants: Olivier Bernard, Antoine Sciandra, Frédéric Grognard, Philipp Hartmann, Ghjuvan Grimaud, Quentin Béchet, David Demory, Hubert Bonnefond, Jean-Philippe Steyer, Francis Mairet.

Experimental developments

Experiments have been carried out to study the effects of nitrogen limitation on the lipid production in microalgae [28] and support model development. These experiments have been carried out in the Lagrangian simulator, under constant or periodic light and temperature, varying the total amount of light dose in the day. The response in terms of storage carbon (triglycerides and carbohydrates) has been observed.

Other experiments were carried out to reproduce the light signal perceived by a cell in a raceway pond [89], derived from hydrodynamical studies [55]. An electronic platform was developed to reproduce this high frequency light signal. The experiments show that the microalgae adapt their pigments to the average light that they have received [28].

The effect in the cell cycle of both the light periodic signal, the temperature and a nitrogen limitation were studied. The strong interactions between the different phases of the cell cycle through checkpoints was highlighted [106]. Temperature turned out to play a key role in modulating metabolic fluxes and synchronization.

The effect of cement flue gas on microalgae growth has been tested. It was demonstrated that this CO_2 source can be used to feed microalgal industrial cultures [114].

Finally a new methodology to measure cell viability has been set up. This approach is very promising to distinguish between net and gross growth rate [22].

These works have been carried out in collaboration with A. Talec, S. Rabouille, E. Pruvost and C. Combe (CNRS/UPMC -Oceanographic Laboratory of Villefranche-sur-Mer).

In collaboration with the IFREMER-PBA team (Nantes) we contributed to a study of the possible associations between microalgae and bacteria to enhance overall productivity [98].

Metabolism of carbon storage and lipid production

A macroscopic model for lipid production by oleaginous microalgae [7] has been previously proposed. This model describes the accumulation of neutral lipids (which can be turned into biofuel), carbohydrates and structural carbon. A metabolic model has been set up and validated for the microalgae *Isochrysis luthea*. It predicts carbohydrate and lipid accumulation, under conditions of light/dark cycles and/or nitrogen deprivation [78], [88], [16].

Modeling the coupling between hydrodynamics and biology

In collaboration with the Inria ANGE team, a model coupling the hydrodynamics of the raceway (based on multilayer Saint-Venant system) with microalgae growth was developed [83]. This model is supported by the work of ANGE aiming at reproducing the hydrodynamics of the raceway, with a specific attention to the effect of the paddle wheel on the fluid [55].

Modeling the photosynthesis response to fast fluctuating light

The impact of the hydrodynamics on the light perceived by a single cell was studied thanks to fluid dynamics simulations of a raceway pond [34]. The light signals that a cell experiences at the Lagrangian scale, depending on the fluid velocity, were then estimated. A Droop-Han model was used to assess the impact of light fluctuation on photosynthesis. A new model accounting for photoacclimation was also proposed [96]. Single cell trajectories were simulated by this software, and the effect on photosynthesis efficiency was assessed using models of photosynthesis [95]. These results were compared to experimental measurements where the high frequency light was reproduced [89].

We also developed a model to reproduce the fluorescence of microalgae during a PAM protocol [51]

Modeling microalgae production processes

The integration of different models developed in the group [81], [101], [7] was performed to represent the dynamics of microalgae growth and lipid production in raceway systems, on the basis of the dynamical model developed to describe microalgal growth in a photobioreactor under light and nitrogen limitations. The strength of this model is that it takes into account the strong interactions between the biological phenomena (effects of light and nitrogen on growth, photoacclimation ...), temperature effect [82], [111] and the radiative transfer in the culture (light attenuation due to the microalgae).
Using these approaches, we have developed a model which predicts lipid production in raceway systems under varying light, nutrients and temperature [109]. This model is used to predict lipid production in the perspective of large scale biofuel production. It was also used to assess the potential of France for microalgae, when taking into account the actual 2012 meteorology at the scale of France the use of lands, slope, proximity of nutrients and CO_2 [73].
In the framework of the ANR project Purple Sun, we develop an innovative system for microalgae production: a raceway pond under a greenhouse with semi-transparent photovoltaic panels. To this end, we include in the microalgae model the effect of light wavelength, and we develop a thermic model of the system in order to estimate the culture temperature.

Finally, we provide guidelines for the design of experiments with high informative content that allows an accurate estimation of the parameters concerning the effect of temperature and light on microalgae growth. The optimal experiment design problem was solved as an optimal control problem. E-optimal experiments were obtained by using two discretization approaches namely sequential and simultaneous. Simulation results showed the relevance of determining optimal experimental inputs for achieving an accurate parameter estimation [39].

Nitrogen fixation by nitrogenotrophs

The fixation of nitrogen by *Croccosphera watsonii* was represented with a macro metabolic model [92]. The main fluxes of carbon and nitrogen are represented in the cell. The accumulation of starch during the day to fuel the nitrogenase working in the absence of oxygen during the night was the key process to explain the nitrogen fixation. The strong influence of the cell cycle was also included in the model. Finally, the model was calibrated and validated with the data of 3 experiments carried out with different duration of the light period and daily dose. The model succeeded to efficiently reproduce the experimental data.

This work is done in collaboration with Sophie Rabouille (CNRS-Oceanographic Laboratory of Villefranche-sur-Mer).

Modeling thermal adaptation in microalgae

We have used the Adaptive Dynamics theory to understand how temperature drives evolution in microalgae. For a constant temperature, we have shown that the optimal temperature trait tends to equal the environment temperature. We then study the case where the temperature is periodically fluctuating [53]. We now use this method at the scale of the global ocean, validating our approach with experimental data sets from 194 species.

Including phytoplankton photoadaptation into biogeochemical models

The complexity of the marine ecosystem models and the representation of biological processes, such as photoadaptation, is very challenging to tackle so that their representation remains an open question. We compared several marine ecosystem models with increasing complexity in the phytoplankton physiology representation in order to assess the consequences of the complexity of photoadaptation models in biogeochemical model predictions. Three models of increasing complexity were considered, and the models were calibrated to reproduce ocean data acquired at the Bermuda Atlantic Time-series Study (BATS) from in situ JGOFS (Joint Global Ocean Flux Study) data. It turns out that the more complex models are trickier to calibrate and that intermediate complexity models, with an adapted calibration procedure, have a better prediction capability [77], [15].

This work is done in collaboration with Sakina Ayata (UPMC-Oceanographic Laboratory of Villefranche-sur-Mer).

6.3.1.2. Control and Optimization of microalgae production

On-line monitoring

Interval observers give an interval estimation of the state variables, provided that intervals for the unknown quantities (initial conditions, parameters, inputs) are known [91]. Several developments were carried out in this direction to improve the design and performances of interval observers, and accounting for a specific structure (*i.e.* triangular) or property (*i.e.* Input to State Stable), [38]. Interval observers were designed for the estimation of the microalgae growth and lipid production within a production process [37] and validated experimentally [36].

Optimization of the bioenergy production systems

Based on simple microalgae models, analytical optimization strategies were proposed. We first focused on the optimal operating conditions for the biomass productivity under day/night cycles using Pontryagin's maximum principle (assuming a periodic working mode) [32].
On the other hand, we assessed strategies for optimal operation in continuous mode using the detailed model for raceways [108], [109]. Two strategies were developed. The first one consists in solving numerically an optimal control problem in which the input flow rate of the raceway is calculated such that the productivity in microalgae biomass is maximized on a finite time horizon. In the second strategy, we translated the optimization problem into a regulation problem. We proposed a simple operational criterion that when integrated in a strategy of closed-loop control allows to attain biomass productivities very near to the maximal productivities obtained with the optimal control. We demonstrated that the practical advantages for real implementation makes our proposed controller a suitable control strategy for optimizing microalgae production in raceways.
We also propose a nonlinear adaptive controller for light-limited microalgae culture, which regulates the light absorption factor (defined by the ratio between the incident light and the light at the bottom of the reactor). We show by numerical simulation that this adaptive controller can be used to obtain near optimal productivity under day-night cycles [103].

Interactions between species

Large scale culture of microalgae for bioenergy involves a large biodiversity (different mutants, invasion, growth-promoting bacteria [98]...). Control of such systems requires to consider the interactions between the different species. Such systems involve bacteria and microalgae, and the competition between these organisms can have several equilibrium points, which can be studied with Monod, Contois and Droop type models [33].

In the framework of the ANR Facteur 4 project, we propose to drive this competition exploring different strategies in order to select species of interest.

We have proposed an adaptive controller which regulates the light at the bottom of the reactor [104]. When applied for a culture with n species, the control law allows the selection of the strain with the maximum growth rate for a given range of light intensity. This is of particular interest for optimizing biomass production as species adapted to high light levels (with low photoinhibition) can be selected.
Strategies to improve the temperature response have been proposed. First we modeled the adaptive dynamics for a population submitted to a variable temperature [53]. This was then used to design experiments aiming at enlarging the thermal niche of a species. Experiments with periodic temperature stresses are currently carried out at the LOV.
Finally, in a more theoretical framework, we studied how to select as fast as possible a given species in a chemostat with two species at the initial instant. Using the Pontryagin maximum principle, we have shown that the optimal strategy is to maintain the substrate concentration to the value maximizing the difference between the growth rates of two species [17]. We now try to extend this result for n species with mutations.

6.3.2. Design of ecologically friendly plant production systems

6.3.2.1. Controlling plant pests

Participants: Frédéric Grognard, Ludovic Mailleret, Suzanne Touzeau, Nicolas Bajeux, Bapan Ghosh.

Optimization of biological control agent introductions

The question of how many and how frequently natural enemies should be introduced into crops to most efficiently fight a pest species is an important issue of integrated pest management. The topic of natural enemies introductions optimization has been investigated for several years [6] [110], unveiling the crucial influence of within-predator density dependent processes. Since parasitoids may be more prone to exhibit positive density dependent dynamics rather than negative ones, which are prevalent among predatory biocontrol agents,the current modeling effort consists in studying the impact of positive predator-predator interactions on the optimal introduction strategies (PhD of Nicolas Bajeux, [45]). The influence of the spatial structure of the environment on biological control efficacy has also been investigated; first results indicate that spatial structure tends to influence it in quite a same way as intra-specific competition does [60]. An extension of that modeling framework was also studied, that considered state dependent impulsive feedback for the stabilization of a positive equilibrium [54].

Connected research on the influence of space on the establishment capacities of biological control agents is also being pursued both through computer simulations and laboratory experiments on parasitoids of the genus *Trichogramma*. This is the topic of the PhD thesis of Thibaut Morel Journel (UMR ISA); in particular, we show how landscape connectivity or spatial heterogeneity shape establishment dynamics in spatially structured environments [63], [64], [65].

Plant compensation, pest control and plant-pest dynamics

Introducing a plant compartment into our models, we first focused on plant-insect interactions and showed how the level and timing of the pest invasion and pests control interventions could have important effects on the plant's growth pattern and its final biomass. We then modeled plant compensation, which is the process by which some plants respond positively to recover from the effects of pest injury. We have shown that depending on plants and pests characteristics, as well as the level of pest attack, plant overcompensation may or may not happen [35]. Experiments have then been held at UMR ISA on tomato plants facing tutta absoluta invasion; tendencies to compensation have been evidenced, but need to be confirmed through larger scale experiments.

This work is part of the PhD thesis of Audrey Lebon (Cirad), supervised in collaboration with Yves Dumont (Cirad), which has been defended in December 2014.

6.3.2.2. Controlling plant pathogens

Participants: Frédéric Grognard, Ludovic Mailleret, Suzanne Touzeau, Elsa Rousseau.

Sustainable management of plant resistance

Because in addition to being eaten, plants can also get sick, we studied other forms of biological control dedicated to fight plant pathogens. One such method is the introduction of plant strains that are resistant to one pathogen. This often leads to the appearance of virulent pathogenic strains that are capable of infecting the resistant plants. It is therefore necessary to find ways to protect the durability of such resistances, which are a natural exhaustible resource. Experiments were conducted in INRA Avignon, followed by high-throughput sequencing (HTS) to identify the dynamics of several virus strains in competition within host plants. Different plant genotypes were chosen for their contrasted effects on genetic drift and selection they induce on virus populations. Those two evolutionary forces can play a substantial role on the durability of plant resistance. Therefore we fitted a mechanistic-statistical model to these HTS data in order to disentangle the relative role of genetic drift and selection during within-host virus evolution [68], [67]. This is the topic of Elsa Rousseau's PhD thesis, and is done in collaboration with Frédéric Fabre and Benoit Moury (INRA Avignon).

We also represented the pathogen spread in agricultural landscapes [40]. At this scale, we looked at how the landscape structure facilitates or impedes the disease spread among host patches. We showed that, when deploying a host with complete resistance to the pathogen along with a susceptible host, mixed landscapes were always more efficient to hamper the disease spread. However, when using a quantitatively resistant host, aggregating the hosts in different regions could result in a better control of the pathogen spread [41]. This work is part of Julien Papaïx's PhD thesis (MIA, INRA Jouy-en-Josas & BIOGER, INRA Grignon).

Eco-evolutionary dynamics of plant pathogens in seasonal environments

Understanding better pathogen evolution also requires to understand how closely related plant parasites may coexist. Indeed, such coexistence is widespread and is hardly explained through resource specialization. We showed that, in agricultural systems in temperate environments, the seasonal character of agrosystems is an important force promoting evolutionary diversification of plant pathogens [94]. Plant parasites reproduction mode may also strongly interact with seasonality. In this context, we investigated the influence of cyclical parthenogenesis, i.e. the alternation of sexual and asexual reproduction phases, on the eco-evolutionary dynamics of plant parasites [25].

This work was part of the PhD thesis of Magda Castel (Agrocampus Ouest) and has been done in collaboration with Frédéric Hamelin (Agrocampus Ouest), Didier Andrivon (INRA Rennes) and Virginie Ravigné (CIRAD Montpellier).

6.3.3. Biological depollution

6.3.3.1. Control and optimization of bioprocesses for depollution

Participants: Olivier Bernard, Francis Mairet, Jean-Luc Gouzé.

We have considered the problem of global stabilization of an unstable bioreactor model (e.g. for anaerobic digestion), when the measurements are discrete and in finite number ("quantized"). These measurements define regions in the state space, wherein a constant dilution rate is applied. We show that this quantized control may lead to global stabilization: trajectories have to follow some transitions between the regions, until the final region where they converge toward the reference equilibrium [71].

Although bioprocesses involve an important biodiversity, the design of bioprocess control laws are generally based on single-species models. In [56], we have proposed to define and study the multispecies robustness of bioprocess control laws: given a control law designed for one species, what happens when two or more species are present? We have illustrated our approach with a control law which regulates substrate concentration using measurement of growth activity. Depending on the properties of the additional species, the control law can lead to the correct objective, but also to an undesired monospecies equilibrium point, coexistence, or even a failure point. We now start to develop control laws more robust to the presence of additional species.

Moreno [107] has proposed an optimal strategy for fed-batch bioreactor with substrate inhibition. Thanks to the Pontryagin maximum principle and the Hamilton-Jacobi equation, we have shown that the same strategy is still optimal when mortality is included in the model [79]. We have also studied the problem when the singular arc is non-necessarily admissible everywhere (i.e. the singular control can take values outside the admissible control set). We have pointed out the existence of a frame point on the singular arc above which any singular trajectory is not globally optimal. Moreover, we have provided an explicit way for computing numerically the switching curves and the frame point [46], [19].

6.3.3.2. Coupling microalgae to anaerobic digestion

Participants: Olivier Bernard, Antoine Sciandra, Jean-Philippe Steyer, Frédéric Grognard, Philipp Hartmann, Francis Mairet.

The coupling between a microalgal pond and an anaerobic digester is a promising alternative for sustainable energy production and wastewater treatment by transforming carbon dioxide into methane using light energy. The ANR Phycover project is aiming at evaluating the potential of this process [113], [112].

In a first stage, we developed models for anaerobic digestion of microalgae. Two approaches were used: first, a dynamic model has been developed trying to keep a low level of complexity so that it can be mathematically tractable for optimization [100] . On the other hand, we have tested the ability of ADM1 [115] (a reference model which considers 19 biochemical reactions) to represent the same dataset. This model, after modification of the hydrolysis step [102] has then been used to evaluate process performances (methane yield, productivity...) and stability though numerical simulations.

Finally, we have proposed and analysed a three dimensional model which represents the coupling of a culture of microalgae limited by light and an anaerobic digester. We first prove the existence and attraction of periodic solutions. Applying Pontryagin's Maximum Principle, we have characterized optimal controls, including the computation of singular controls, in order to maximize methane production. Finally, we determine numerically optimal trajectories by direct and indirect methods [18].

6.3.3.3. Life Cycle Assessment

Participants: Olivier Bernard, Jean-Philippe Steyer.

This work is the result of a collaboration with Arnaud Helias of INRA-LBE and Pierre Collet (IFPEN).

In the sequel of the pioneering life cycle assessment (LCA) work of [97], we continued to identify the obstacles and limitations which should receive specific research efforts to make microalgae production environmentally sustainable.

The improvements due to technological breakthrough (leading to higher productivities) have been compared to the source of electricity. It turns out that the overall environmental balance can much more easily be improved when renewable electricity is produced on the plant [27]. As a consequence, a new paradigm to transform solar energy (in the large) into transportation biofuel is proposed, including a simultaneous energy production stage. This motivated the design of the purple sun ANR-project where electricity is produced by semi transparent photovoltaic panels [74] under which microalgae are growing.

These studies have allowed to identify the obstacles and limitations which should receive specific research efforts to make this process environmentally sustainable [93].

Finally, some works are aiming at normalizing LCA for microalgae and proposing guidelines to make the LCA more easily comparable [87].

These works have been carried out in collaboration with E. Latrille and B. Sialve (INRA - Laboratory of Environmental Biotechnology, Narbonne).

6.3.4. Models of ecosystems

6.3.4.1. Optimality/games in population dynamics

Participants: Frédéric Grognard, Ludovic Mailleret, Pierre Bernhard.

Optimal foraging and residence times variations

In a pair of papers [23], [24], we reanalyzed the so-called Marginal Value Theorem (MVT), first published in 1976. This theorem, also used in human behavior and economics, establishes how individuals should behave to optimize resource exploitation. This result has been has been routinely applied in ecology to understand the foraging strategy of animals such as insect parasitoids used for biological control purposes. We obtained the first mathematical characterization of how habitat characteristics (e.g. patch quality, or the distance between resource patches) affect the optimal foraging strategy. This allowed to confirm or refine MVT predictions, and to provide new predictions in the more realistic case of heterogeneous habitats. Some counterintuitive predictions emerged: making resource patches richer can actually make individuals move more rapidly, contradicting generally admitted earlier predictions.

This work was conducted with Vincent Calcagno (UMR ISA) and Frédéric Hamelin (Agrocampus Ouest).

The handicap paradox

We have continued our investigation of the handicap paradox of sexual selection with the tools of signaling theory. Zahavi's handicap principle, and our game theoretic analysis, explain why an equilibrium displays the "handicap" feature [21]. However, the explanation seems somewhat contrived, so the next question is "how could evolution have reached such a state ?" We have investigated that question with the tools of adaptive dynamics, and reached the conclusion that, if one accepts adaptive dynamics as a model of evolution, and our model of sexual selection, the equilibrium described in our previous article is indeed the limit state of evolution [20].

This work was conducted with Frédéric Hamelin (Agrocampus Ouest).

BIPOP Project-Team

6. New Results

6.1. Highlights of the Year

- Bernard Brogliato: keynote speaker at ICDVC-2014, 4th International Conference on Dynamics, Vibration and Control, August 23-25, 2014 in Shanghai, China. http://www.icdvc2014.org/

6.2. Multiple impacts modelling

Participant: Bernard Brogliato.

The work consists of studying two systems: the rocking block and tapered chains of balls, using the Darboux-Keller model of multiple impacts previously developed. The objectives are threefold: 1) show that the model predicts well the motion by careful comparisons with experimental data found in the literature, 2) study the system's dynamics and extract critical kinetic angles that allow the engineer to predict the system's gross motion, 3) develop numerical code inside the SICONOS platform that incorporates the model of multiple impact. The influence of the kinetic angles in the rocking block motion with friction is analysed as well, numerically. Extensive experimental works have been conducted by our colleague C. Liu at PKU on a disc-ball system. Results are in [32] [67], and in the monograph [16]. Multiple impacts have also been tackled through generalized kinematic models using the kinetic metric [20].

6.3. The contact complementarity problem

Participants: Bernard Brogliato, Florence Bertails-Descoubes, Alejandro Blumentals.

The contact linear complementarity problem is an set of equalities and complementarity conditions whose unknowns are the acceleration and the contact forces. It has been studied in a frictionless context with possibly singular mass matrix and redundant constraints in [21], using results on well-posedness of variational inequalities obtained earlier by the authors. This is also the topic of the first part of the Ph.D. thesis of Alejandro Blumentals where the frictional case is treated as a perturbation of the frictionless case. The contact LCP is directly related to the so-called Painlevé's paradox of contact mechanics. In collaboration with C. Liu (Beijing university PKU) some results habe been obtained from the analysis of a compliant model in the limit. It shows on the classical sliding rod system that the inconsistent mode yield to instantaneous transition to a sticking mode. This is quite coherent with previous results obtained by Le xuan Anh in 1991 on the Painlevé-Klein system (bilateral constraints with Coulomb friction). The results will appear in Multibody System Dynamics in 2015.

6.4. Discrete-time sliding mode control

Participants: Vincent Acary, Bernard Brogliato, Olivier Huber.

This topic concerns the study of time-discretized sliding-mode controllers. Inspired by the discretization of nonsmooth mechanical systems, we propose implicit discretizations of discontinuous, set-valued controllers. This is shown to result in preservation of essential properties like simplicity of the parameters tuning, suppression of numerical chattering, reachability of the sliding surface after a finite number of steps, and disturbance attenuation by a factor h or h^2 [41], [42], [43], [45], [61]. This work is part of the ANR project CHASLIM. Within the framework of CHASLIM we have performed many experimental validations on the electropneumatic setup of IRCCyN (Nantes), which nicely confirm our theoretical and numerical predictions: the implicit implementation of sliding mode control, drastically improves the input and output chattering behaviours. In particular the high frequency bang-bang controllers which are observed with explicit discretizations, are completely suppressed. The implicit discretization has been applied to the classical equivament-based-control SMC, and also to the twisting sliding-mode controller [43].

6.5. Lur'e set-valued dynamical systems

Participants: Bernard Brogliato, Christophe Prieur.

Lur'e systems are quite popular in Automatic Control since the fifties. Set-valued Lur'e systems possess a static feedback nonlinearity that is a multivalued function. This study consists in the mathematical analysis (existence and uniqueness of solutions) and the stability analysis (Lyapunov stability, invariance principle) of classes of set-valued Lur'e systems, with applications in complementarity dynamical systems, relay systems, mechanical systems with dry friction, electrical circuits, etc. Our works in this field started in [62]. The results in [64] extend those in [63] with an accurate characterization of the maximal monotonicity of the central operator of these systems, which consists of a projection-like operator. Concrete and verifiable criteria are provided for the above classes (complementarity, relay systems). Results on state observers and output feedback control for classes of Lur'e systems (namely: Moreau's sweeping process of first and second order, and with prox-regular sets) are proposed in [29], [44], [34]. Therein the convexity is replaced by the far more general notion of prox-regularity, which destroys the monotonicity. The input to state stability of measure driven differential equations has been tackled in [22], where some results from [29] are adapted.

6.6. Simulation and stability of piecewise linear gene networks

Participants: Vincent Acary, Arnaud Tonnelier, Bernard Brogliato.

This work has been done in collaboration with the IBIS project team, it is reported in [19]. Gene regulatory networks control the response of living cells to changes in their environment. A class of piecewise-linear (PWL) models, which capture the switch-like interactions between genes by means of step functions, has been found useful for describing the dynamics of gene regulatory networks. The step functions lead to discontinuities in the right-hand side of the differential equations. This has motivated extensions of the PWL models based on differential inclusions and Filippov solutions, whose analysis requires sophisticated numerical tools. We present a method for the numerical analysis of one proposed extension, called Aizerman-Pyatnitskii (AP)-extension, by reformulating the PWL models as a mixed complementarity system (MCS). This allows the application of powerful methods developed for this class of nonsmooth dynamical systems, in particular those implemented in the Siconos platform. We also show that under a set of reasonable biological assumptions, putting constraints on the right-hand side of the PWL models, AP-extensions and classical Filippov (F)-extensions are equivalent. This means that the proposed numerical method is valid for a range of different solution concepts. We illustrate the practical interest of our approach through the numerical analysis of three well-known networks developed in the field of synthetic biology.

In addition, we have investigated oscillatory regimes in repressilator-type models with piecewise linear dynamics [30]. We derived exact analytical conditions for oscillations and showed that the relative location between the dissociation constants of the Hill functions and the ratio of kinetic parameters determines the possibility of oscillatory activities. We also computed analytically the probability of oscillations. Results suggest that a switch-like coupling behaviour, a time-scale separation and a repressilator-type architecture with an even number of elements facilitate the emergence of sustained oscillations in biological systems.

6.7. Numerical analysis and simulation of mechanical systems with constraints

6.7.1. Event-capturing schemes for nonsmooth mechanical systems

Participant: Vincent Acary.

To perform the numerical time integration of nonsmooth mechanical systems, the family of event-capturing time–stepping schemes are the most robust and efficient tools. Nevertheless, they suffer from several draw-backs : a) a low–order accuracy (at best at order one), b) a drift phenomena when the unilateral constraints are treated at the velocity level and c) a poor "energetic" behavior in terms of stabilizing the high–frequency dynamics. We proposed self-adapting schemes by applying time–discontinuous Galerkin methods to the mea-sure differential equation in [28]. In order to satisfy in discrete time, the impact law and the constraints at the position and the velocity level, an adaptation of the well-known Gear–Gupta–Leimkuhler approach has been developed. In [58], the approach is algorithmically specified, improved and applied to nonlinear multi-contact examples with friction. It is compared to other numerical schemes and it is shown that the newly proposed integration scheme yields a unified behavior for the description of contact mechanical problems. Especially, we provide time-integration of the nonimpulsive dynamics with semi–explicit Runge–Kutta method previously developed for differential algebraic equations.

6.7.2. Numerical time–integration methods for event-detecting schemes.

Participants: Vincent Acary, Bernard Brogliato, Mounia Haddouni.

The CIFRE thesis of M. Haddouni concerns the numerical simulation of mechanical systems subject to holonomic bilateral constraints, unilateral constraints and impacts. This work is performed in collaboration with ANSYS and the main goal is to improve the numerical time–integration in the framework of event-detecting schemes. Between nonsmooth events, time integration amounts to numerically solving a differential algebraic equations (DAE) of index 3. We have compared dedicated solvers (Explicit RK schemes, Half-explicit schemes, generalizes α-schemes) that solve reduced index formulations of these systems. Since the drift of the constraints is crucial for the robustness of the simulation through the evaluation of the index sets of active contacts, we have proposed some recommendations on the use of the solvers of dedicated to index-2 DAE. A manuscript has been submitted to Multibody System Dynamics.

6.7.3. Multibody systems with clearances (dynamic backlash)

Participants: Vincent Acary, Bernard Brogliato, Narendra Akadkhar.

The PhD thesis of N. Akadkhar under contract with Schneider Electric concerns the numerical simulation of mechanical systems with unilateral constraints and friction, where the presence of clearances in imperfect joints plays a crucial role. A first work deals with four-bar planar mechanisms with clearances at the joints, which induce unilateral constraints and impacts, rendering the dynamics nonsmooth. The objective is to determine sets of parameters (clearance value, restitution coefficients, friction coefficients) such that the system's trajectories stay in a neighborhood of the ideal mechanism (*i.e.* without clearance) trajectories. The analysis is based on numerical simulations obtained with the projected Moreau-Jean time-stepping scheme. These results have been reported in [37]. It is planned to extend these simulations to frictional cases and to mechanisms of circuit breakers.

6.8. Inverse modeling with contact and friction

6.8.1. Inverse statics of plates and shells

Participants: Florence Bertails-Descoubes, Romain Casati, Gilles Daviet.

We have started to investigate the problem of interpreting an arbitrary 3D mesh as an equilibrium configuration of an elastic plate/shell, subject to gravity and frictional contact forces. We have first considered a simple nodal shell model accounting for stretch, shear and bending. For such a model, inverse statics formulates as an ill-posed minimization problem with a nonlinear objective and nonsmooth constraints. Our objective is to examine this problem in the case where the rest pose of the system is left as unknown, while material parameters (mass, stiffness) are assumed to be known (inverse design problem). In some specific cases (cloth modeling), we use a priori information such as locally low Gaussian curvature so as to help the retrieval of most natural solutions. We plan to submit our results to Siggraph 2015. Targeted applications include virtual garment modeling and parameter retrieval from 3D image-based capture.

6.9. Modeling of fibrous medium

6.9.1. Continuous modeling of fiber assemblies

Participants: Florence Bertails-Descoubes, Gilles Daviet.

Following the exploratory project funded by Persyval (2013-2014), we have started to model an assembly of long elastic fibers (such as hair) using a continuous approach (continuum mechanics equations coupled with a nonsmooth stress-strain law). Interactions between air and fibers can then be naturally accounted for, increasing the realism of some macroscopic features compared to our previous discrete elements model. This is still work in progress and we will make some of our results publicly available in 2015.

6.10. Threshold in spiking neural models

Participant: Arnaud Tonnelier.

We studied the threshold for spike initiation in two-dimensionnal spiking neural models. A threshold criterion that depends on both the membrane voltage and the recovery variable is proposed. This approach provides a simple and unified framework that accounts for numerous voltage threshold properties. Implications for neural modeling are also discussed [31].

6.11. Nonsmooth modes in chains of impact oscillators

Participants: Vincent Acary, Guillaume James, Franck Pérignon.

Chains of impact oscillators arise for example as finite-element models of thin oscillating mechanical structures (a string under tension or a clamped beam) contacting rigid obstacles. Nonlinear periodic waves are observed in experiments on such systems, but relatively little is known from a theoretical point of view on their existence and stability. In 2008, Gendelman and Manevitch have analyzed the existence and stability of nonlinear localized modes (breathers) for discrete linear chains with a single node undergoing rigid impacts. In this work, we introduce a numerical method allowing to compute branches of time-periodic solutions when an arbitrary number of nodes undergo rigid impacts without energy dissipation. For this purpose, we reformulate the search of periodic solutions as a boundary value problem incorporating unilateral constraints. We illustrate this numerical approach by computing different families of breathers and nonlinear normal modes. Our method is much more effective than a numerical continuation of periodic solutions based on compliant models, which requires to integrate stiff differential equations and lead to costly numerical continuation. These results have been communicated in two international conferences, ENOC 2014 [35] and 11th World Congress on Computational Mechanics [36].

6.12. Traveling waves in spatially discrete excitable media

Participants: José Eduardo Morales, Arnaud Tonnelier, Guillaume James.

The propagation of traveling waves in excitable media is a widespread phenomenon, with applications ranging from forest fires to electrical signals propagating along nerve fibers. The case of spatially discrete excitable models is notoriously difficult to analyze. In particular, for the discrete FitzHugh-Nagumo reaction-diffusion system, the existence of pulses for a general class of bistable nonlinearities has been proved only recently (Hupkes and Sandstede, 2010). The existence of pulses under more general types of interactions (e.g. elastic instead of diffusive) remains an open question, as well as traveling wave propagation in higher-dimensional systems. These problems will be tackled in the PhD thesis of J.-E. Morales (advisors A. Tonnelier and G. James), which started on November 2013. J.-E. Morales has started to analyze pulse propagation in the excitable Burridge-Knopoff (BK) model, which finds applications in the context of nonlinear friction. This model includes elastic interactions between particles, and an additional difficulty linked with nonsmoothness of the (multivalued) Coulomb friction law. Using an idealized piecewise linear friction force, we have studied the propagation of a pulse wave in the discrete BK model. Using asymptotic methods, we proved the existence of a pulse wave and derived quantitative results for travelling wave properties.

6.13. Nonlinear waves in granular chains

Participants: Guillaume James, Bernard Brogliato.

Granular chains made of aligned beads interacting by contact (e.g. Newton's cradle) are widely studied in the context of impact dynamics and acoustic metamaterials. When a large number of beads are present, their dynamics can be described by infinite-dimensional differential equations, which possess a limited smoothness when unilateral Hertzian contact interactions are considered. In this context, we have developed and analyzed new reduced-order models describing nonlinear wave propagation in such systems. In the work [25] (collaboration with D.Pelinovsky, McMaster Univ.), we analyze small amplitude slowly modulated compression waves in the limit when the exponent of the Hertz force is close to unity. From a multiple scale analysis, we derive a Korteweg-de Vries equation with logarithmic nonlinearity allowing to approximate wave profiles, in particular solitary wave solutions. In the work [50] (collaboration with Y. Starosvetsky, Technion IIT), we prove existence of spatially localized nonlinear modes (breathers) in the DpS equation, an amplitude equation describing small oscillations in Newton's cradle over long time scales. For Hertz force exponents close to unity, we show that breather envelopes are well approximated by a Gaussian solution of the logarithmic nonlinear Schrödinger equation. This result is generalized to traveling localized oscillations (traveling breathers) generated by an impact in Newton's cradle (G. James, article in preparation). The existence of breathers is also analyzed in granular metamaterials consisting of hollow beads with internal masses (G. James) in collaboration with L. Liu, A. Vainchtein (Pittsburgh Univ.) and P. Kevrekidis (UMass Amherst) - article in preparation. In addition the LZB model introduced in [15] has been extensively used to numerically investigate wave phenomena in chains of aligned balls (tapered, monodisperse, anti-tapered, stepped chains). Thorough comparisons with experimental results reported in the Granular Matter literature have been made. The results are reported in the monograph [16].

6.14. Robotics

6.14.1. Lexicographic Least-Squares solver

Participants: Pierre-Brice Wieber, Dimitar Dimitrov.

We have been working on Multi-Objective Least-Squares problems with inequality constraints for the last few years, focusing especially on the Lexicographic case. The focus this year has been on nonlinear problems, in collaboration with Adrien Escande from JRL, Tsukuba, Japan. Questions concerning the second-order approximation, using a Gauss-Newton approach or considering more precise second-order information, and questions concerning the globalization scheme, trust-region and/or filter methods have been approached, but results are still preliminary.

6.14.2. Mobile manipulation by humanoid robots

Participants: Pierre-Brice Wieber, Dimitar Dimitrov, Alexander Sherikov, Jory Lafaye.

The realization of mobile manipulation by humanoid robots requires the handling of two simultaneous problems: taking care of the dynamic balance of the robot, what is usually done with Model Predictive Control (MPC) schemes, and redundant motion and force control of the whole body of the robot, what is usually done with a Quadratic Program, or a more advanced Lexicographic Least-Squares problem (see above). These two problems are usually solved in sequence: an MPC scheme first computes the necessary motion of the feet and Center of Mass (CoM) of the robot, then motion and force redundancy of the whole body of the robot is resolved. We have observed that this sequence corresponds to a lexicographic order between two objectives, feet and CoM motion first, the rest of the body after, which limits the possibility to tackle scenarios where we would like the motion of the CoM of the robot to be driven by the motion of the rest of the body of the robot, for example to catch an object with the hand. We have proposed therefore to reorganize the order between these different objectives, building on the LexLS solver presented above. The focus this year has been on non-coplanar multi-contact situations.

6.14.3. Reactive trajectory generation

Participants: Pierre-Brice Wieber, Dimitar Dimitrov, Saed Al Homsy, Matthieu Guilbert.

The goal of the ongoing collaboration with Adept Technologies is to generate near time optimal trajectories in the presence of moving obstacles in real time. Results are not public yet due to industrial constraints.

6.15. Optimization

6.15.1. Semidefinite programming and combinatorial optimization

Participant: Jérôme Malick.

We have worked with Frederic Roupin (Prof. at Paris XIII) and Nathan Krislock (Assistant Prof. at North Illinois University, USA) on the use of semidefinite programming to solve combinatorial optimization problems to optimality.

We proposed a new family of semidefinite bounds for 0-1 quadratic problems with linear or quadratic constraints [65]. We have embedded the new bounds within branch-and-bound algorithms to solve 2 standard combinatorial optimization problems to optimality.

- *Max-cut.* We developed [26] an improved bounding procedure obtained by reducing two key parameters (the target level of accuracy and the stopping tolerance of the inner Quasi-Newton engine) to zero, and iteratively adding triangle inequality cuts. We also precisely analyzed its theoretical convergence properties. We show that our method outperform the state-of-the-art solver ([66]) on the large test-problems.
- *Heaviest k-subgraph problems*. Adapting the techniques we developped for the max-cut problem, we have proposed in [60] an algorithm able to solve exactly k-cluster instances of size 160. In practice, our method works particularly fine on the most difficult instances (with a large number of vertices, small density and small k).

We have also been working on a generic online semidefinite-based solver for binary quadratic problems using the generality of [65]. Finally, a first web interface for our solvers and our data sets are available online at http://lipn.univ-paris13.fr/BiqCrunch/.

6.15.2. Quadratic stabilization of Benders decomposition

Participants: Jérôme Malick, Sofia Zaourar.

The Benders decomposition, a fundamental method in operation research, is known to have the inherent instability of cutting plane-based methods. The PhD thesis of Sofia Zaourar proposes a algorithmic improvement of the method inspired from the level-bundle methods of nonsmooth optimisation. We illustrate the interest of the stabilization on two classical network problems: network design problems and hub location problems. We also prove that the stabilized Benders method have the same theoretical convergence properties as the usual Benders method. An article about this research was submitted this summer.

BONSAI Project-Team

6. New Results

6.1. Highlights of the Year

- Amandine Perrin received the best paper award and the best oral presentation at the ISCB-LA 2014 international conference for the work on reconstruction of ancestral gene orders.
- Hélène Touzet was invited as a keynote speaker at the ALGO 2014 international conference. The topic of the talk was RNA bioinformatics.

BEST PAPERS AWARDS :

[7] **ISCB-Latin America**. A. PERRIN, J.-S. VARRÉ, S. BLANQUART, A. OUANGRAOUA.

6.2. High-throughtput sequence processing

- **Analysis of immunological rearrangements for leukemia diagnosis and monitoring.** High-throughput sequencing is spreading in the hospitals and many classical routines are now being transferred to this new technology. However in the specific case of lymphocyte monitoring, some complications arise. Classical bioinformatics software tools do not apply to the specificity of lymphocyte rearrangements. That is why we developed the software Vidjil (see 5.2) together with Lille hospital. This work has been published [5] and was also presented, as a poster, during the annual conference of the American Society of Hematology (ASH) [13]. We are now members of the EuroClonality-NGS work group which aims at providing a standardized way of monitoring leukemia using high-throughput sequencing at the European level.
- **New seeds for approximate pattern matching.** We addressed the problem of approximate pattern matching using the Levenshtein distance. Given a text T and a pattern P, find all locations in T that differ by at most k errors from P. For that purpose, we proposed a filtration algorithm that is based on a novel type of seeds, combining exact parts and parts with a fixed number of errors, that we called 01^*0 seed. Implementation has been performed on a Burrows-Wheeler transform. Experimental tests show that the method is specifically well-suited to search for short patterns (< 50 letters) on a small alphabet (*e.g.* DNA alphabet) with a medium to high error-rate (7 %–15 %). This work has been published in [9], and has a large number of applications in computational biology, such as finding microRNA targets, for example.
- **Spaced seeds and Transition seeds.** This year, two collaborative works have been published on the topic of spaced seeds and derivated models. The first work, resulting from a collaboration with Martin C. Frith from the *Computational Biology Research Center* (Tokyo), increases the sensivity of several search tools (among them, LAST, LASTZ, YASS,...) by computing specific seeds adapted to transition ratios observed during Eucaryotic comparisons. This work has been published in [3], together with the design of seeds obtained. The second work, issued from collaboration with Donald E.K. Martin from the *Department of Statistics* of the *North Carolina State University* (Raleigh), deals with the coverage of spaced seeds and shows how this criterion helps selecting good seeds for SVM string-kernels and alignment-free distances. This work has been published in [6].

6.3. RNA algorithms

- **A universal framework for RNA algorithms.** We have proposed a new generic specification framework, called *inverted coupled rewrite systems* that can deal with optimization problems on strings, trees, and arc-annotated sequences. It is specifically well-suited to handle RNA algorithms, such as alignment or folding algorithms. It is based on the following ideas. The solutions of combinatorial optimization problems are the inverse image of a term rewrite relation that reduces problem solutions to problem inputs. A tree grammar is used to further refine the search space, and optimization objectives are specified as interpretations of these terms. All these constituents provide a mathematically precise and complete problem specification, leading to concise yet translucent specifications of dynamic programming algorithms. This work is a collaborative project with R. Giegerich from Universität Bielefeld, and has been published in [4].

- **RNA multistructures.** In many RNA families, the signature of the family cannot be characterized by a single consensus structure, and is mainly described by a set of alternate secondary structures. For example, certain classes of RNAs adopt at least two distinct stable folding states to carry out their function. This is the case of riboswitches, that undergo structural changes upon binding with other molecules, and recently some other RNA regulators were proven to show evolutionary evidence for alternative structure. The necessity to take into account multiple structures also arises when modeling an RNA family with some structural variation across species, or when it comes to work with a set of predicted suboptimal foldings. In this perspective, we have introduced the concept of RNA multistructures, that is a formal grammar based framework specifically designed to model a set of alternate RNA secondary structures. We provide several motivating examples and propose an efficient algorithm to search for RNA multistructures within a genomic sequence. This work was published in [8].

6.4. Ancestral gene order reconstruction

- In the field of **genomic rearrangement**, a topic of interest is to infer ancestral gene order from gene order known in extant species. The problem resumes to compute a set ancestral CARs (continuous ancestral regions) at a given node of a phylogeny. We designed a progressive homology-based method which iteratively detects and assembles ancestral adjacencies while allowing some micro-rearrangements of synteny blocks at the extremities of the progressively assembled CARs. Comparing with other methods we are able to produce more robust CARs with a very simple and efficient method. This work was published in [7].

6.5. Nonribosomal peptides

- **Monomeric structure.** The algorithm that identifies the monomeric structure of a polymer from its chemical structure has been finished and named s2m. It is based on a double index: A partial index constructed on the monomer database that uses a markovian model to speed up the search time ; and an index constructed on the fly on the studied polymer. This strategy was originally developed for nonribosomal peptides, but can be applied to any polymer.
- **Florine: Nonribosomal peptide synthetase annotations.** Florine [2] is a workflow dedicated to the discovery of new nonribosomal peptide synthetases. It describes sequential steps starting from DNA sequences leading to the design of candidate bioactive peptides. It is a useful tool for new drug discovery because it can be applied whatever the producing micro-organisms as it takes into account the enzymatic specificities related to each genus. This work was performed in collaboration with members of EPI Orpailleur (CRI Nancy Grand Est), Marie-Dominique Devignes and Malika Smaïl-Tabbone.
- **Activity prediction of small molecules.** Bayesian Belief Network was used for the first time to classify compounds according to their biological activity [1]. This method was applied on nonribosomal peptides and gave promising results on predicting their activity.

CAGIRE Team

6. New Results

6.1. DNS of a Taylor Green vortex

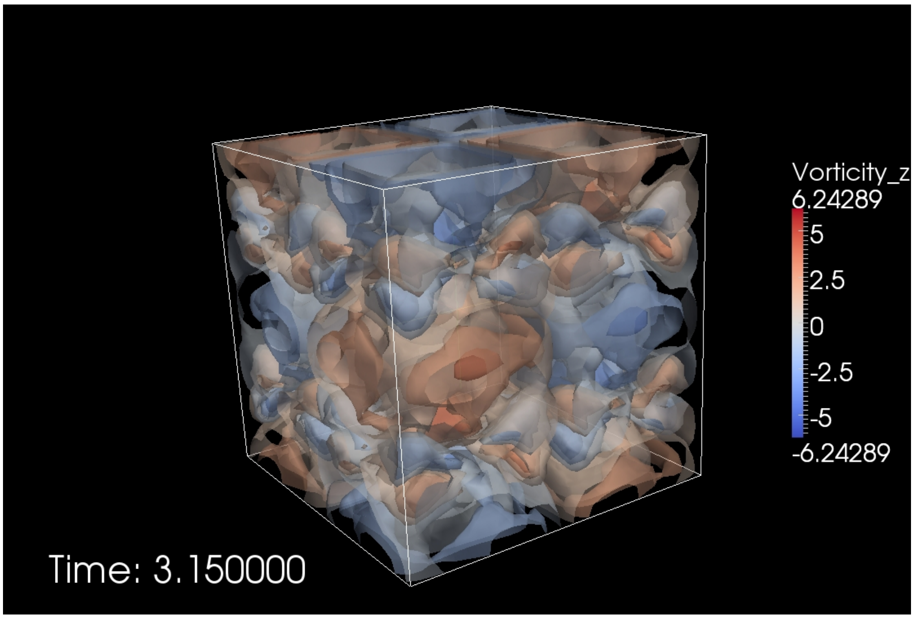


Figure 3. AeroSol simulation of a Taylor-Green vortex: snapshot of one component of the vorticity.

In 2014, we finished the validation of Navier-Stokes discretization with the discontinuous Galerkin method in the Aerosol library. The result of Figure 3 is a first validation in turbulence conditions. The Taylor-Green vortex case is part of the C3 (i.e. "difficult") test cases of the high order CFD workshop, see https://www.grc.nasa.gov/hiocfd/.

6.2. Low Mach number flows simulations issues

Our activity for developing schemes suitable for the simulation of low Mach number flows considers the two main techniques developed initially for dealing with either zero Mach number flows (pressure-velocity coupling) or compressible flows (density based approach). For both approaches, we concentrated this year on the specific difficulties related to unsteady flows simulations. For the methodology adressing the pressure-velocity coupling with a low-order discretization technique, we introduced an inertia term in the AUSM+ -up scheme. The resulting scheme, called AUSM-IT (IT for Inertia Term), was designed as an extension of the AUSM+ -up scheme allowing for full Mach number range calculations of unsteady flows including acoustic features. In line with the continuous asymptotic analysis, the AUSM-IT scheme satisfies the conservation of the discrete linear acoustic energy at first order in the low Mach number limit. Its capability to properly handle low Mach number unsteady flows, that may include acoustic waves or discontinuities was numerically illustrated [7].

As far as density based approach are concerned, an analysis of explicit RKDG schemes have been performed for unstationary acoustic waves propagating in a low Mach number flow. Classical cures of the unaccuracy of upwind schemes at low Mach number consist in using centered flux on the pressure. By a two scale asymptotic expansion of the scheme, we proved that this cure is a dead end for resolving unstationary acoustic waves, because it leads to a non dissipative scheme for the wave equations. We developed a dissipative term that can both stabilize the stationary incompressible equations, and the system of acoustic waves. The results with this new type of scheme have been presented in [8].

6.3. Improving the flexibility of turbulence models for industrial applications

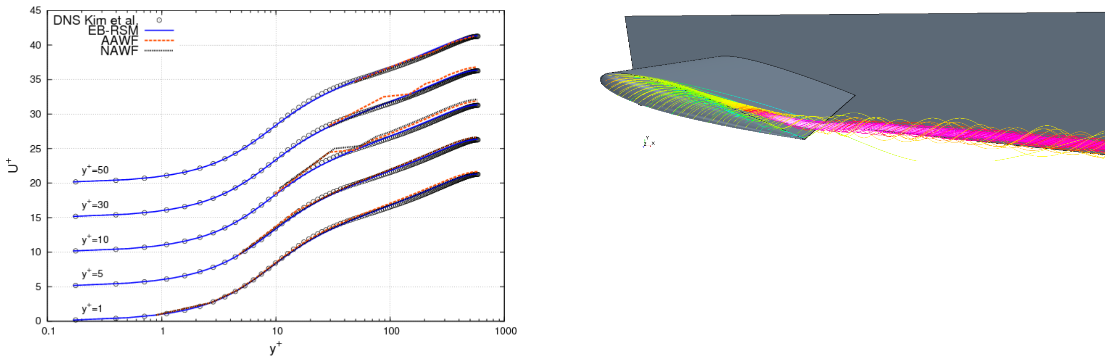


Figure 4. Left: Computation (Code_Saturne) of turbulent channel flow at 3 Reynolds numbers. Comparison with reference DNS of the results given by the EB-RSM integrated down to the wall (ItW, fine mesh) and the EB-RSM with analytical adaptive wall function (AAWF, 3 meshes). Right: EB-RSM computation (STARCCM+ code) of the wing-tip vortex generated by the flow around a NACA 0012 at 10 deg incidence. Visualisation of the streamlines colored with the streamwise vorticity.

In collaboration with industrial partners (EDF and CD-Adapco) developing CFD codes (code_Saturne and STARCCM+, respectively), we are working on the flexibility and robustness of the EB-RSM, an advanced Reynolds-stress turbulence model. Indeed, the two main problems that slow down the spreading of the use of such low-Reynolds number models (i.e., integrating the equations down to solid boundaries) in the industry are the impossiblity to control the near-wall mesh quality in the whole domain of a complex industrial application and the occurence of numerical instabilities due to spurious relaminarizations in some configurations.

In order to address the first issue, we are working, in particular in the frame of the PhD thesis of J.-F. Wald, on the development of adaptive wall functions, i.e., non-homogeneous Dirichlet boundary conditions for the turbulent variables dependant on the size of the cell adjacent to the wall. These wall functions are based on the physical properties of turbulence in the different layers of the near-wall region (asymptotic behaviour in the viscous sublayer and log law in the equilibrium layer), such a way that the flow is correctly reproduced whatever the near-wall refinement of the mesh. Fig. 4 (left) shows that the reproduction of the mean velocity profile in turbulent channel flows obtained using a typical, industrial mesh ($y^+ = 50$) remains very close to the grid-converged solution.

The second issue, the numerical instabilities due to local, spurious relaminarization of the model, can be addressed by investigating the solutions of the dynamical system formed by the model equations in homogeneous situations. Equilibrium solution are intersections of the nullclines (the locus of steady solutions for individual equations) and the stability properties of these fixed points can be visualized using trajectories in the phase space. By investigating the dependance of these stability properties on the parameters of the model, it is possible to eliminate undesired stable fixed points and thus to avoid the appearance of spurious laminarization. Fig. 4 (right) shows the fully turbulent solution obtained with the modified model in a case where the original model exhibited a severe, unphysical relaminarization of the wing-tip vortex.

6.4. Assessment of the discontinuous Galerkin methods on curved meshes

The internship of Hamza Belkhayat-Zougari was concerned with the handling of high order curved meshes in the Aerosol library. During his internship, we developed new analytical solutions of the Laplace and of the Navier-Stokes equations on curved domains for emphasizing the limitation at second order of high order methods on straight meshes, and for assessing the right order on high order meshes. Example of order obtained on straight and curved meshes can be found on Figure 5 .

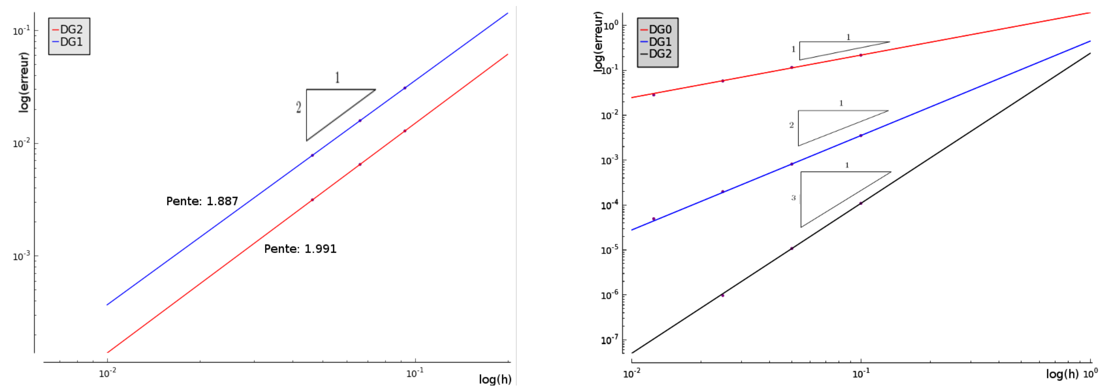


Figure 5. Convergence on a ring for the Laplace equation. Left: high order method on a straight mesh is limited to two. Right: third order accuracy can be recovered by using a second order mesh.

CAIRN Project-Team

6. New Results

6.1. Highlights of the Year

Our work on accuracy evaluation and optimisation for fixed point arithmetic was presented during a tutorial "Automatic Fixed-Point Conversion: a Gateway to High-Level Power Optimization" at IEEE/ACM Design Automation and Test in Europe [77].
As a proof of concept of our research on improving efficiency of dynamic reconfiguration in FPGAs [47] [48], the *eFPGA* (Figure 5) chip was designed and fabricated in 65nm CMOS technology. In the proposed and patented architecture [73] (EU patent), the configuration of the FPGA becomes independent from its placement and is moreover significantly compressed (up to $\times 10$). This notion of *Virtual Bit Stream* allows for seamless partial and dynamic reconfiguration and for task migration.

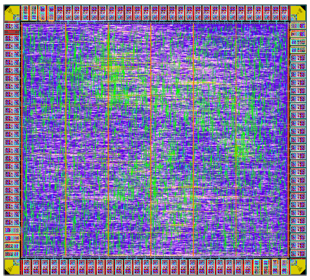


Figure 5. CAIRN*'s eFPGA chip*

6.2. Reconfigurable Architecture Design

6.2.1. Dynamic reconfiguration support in FPGA

Participants: Olivier Sentieys, Antoine Courtay, Christophe Huriaux.

Almost since the creation of the first SRAM-based FPGAs there has been a desire to explore the benefits of partially reconfiguring a portion of an FPGA at run-time while the remainder of design functionality continues to operate uninterrupted. Currently, the use of partial reconfiguration imposes significant limitations on the FPGA design: reconfiguration regions must be constrained to certain shapes and sizes and, in many cases, bitstreams must be precompiled before application execution depending on the precise region of the placement in the fabric. We plan to develop an FPGA architecture that allows for seamless translation of partially-reconfigurable regions, even if the relative placement of fixed-function blocks within the region is changed.

FPGA Architecture Support for Heterogeneous, Relocatable Partial Bitstreams.

The use of partial dynamic reconfiguration in FPGA-based systems has grown in recent years as the spectrum of applications which use this feature has increased. For these systems, it is desirable to create a series of partial bitstreams which represent tasks which can be located in multiple regions in the FPGA fabric. While the transferal of homogeneous collections of lookup-table based logic blocks from region to region has been shown to be relatively straightforward, it is more difficult to transfer partial bitstreams which contain fixed-function resources, such as block RAMs and DSP blocks. In this work we consider FPGA architecture enhancements which allow for the migration of partial bitstreams including fixed-function resources from region to region even if these resources are not located in the same position in each region. Our approach does not require significant, time-consuming place-and-route during the migration process. We quantify the cost of inserting additional routing resources into the FPGA architecture to allow for easy migration of heterogeneous, fixed-function resources. Our experiments show that this flexibility can be added for a relatively low overhead and performance penalty. This work was performed during Christophe Huriaux's visit at UMASS in summer 2014 in the context of Inria Associate Team Hardiesse and has been published in [48] and in [74] as a poster.

Virtual Bit Streams: Design Flow and Run-Time Management of Compressed and Relocatable FPGA Configurations.

The aim of partially and dynamically reconfigurable hardware is to provide an increased flexibility through the load of multiple applications on the same reconfigurable fabric at the same time. However, a configuration bit-stream loaded at runtime should be created offline for each task of the application. Moreover, modern applications use a lot of specialized hardware blocks to perform complex operations, which tends to cancel the "single bit-stream for a single application" paradigm, as the logic content for different locations of the reconfigurable fabric may be different. We proposed a design flow for generating compressed configuration bit-streams abstracted from their final position on the logic fabric. Those configurations can then be decoded and finalized in real-time and at run-time by a dedicated reconfiguration controller to be placed at a given physical location. The VTR framework has been expanded to include bit-stream generation features. A bit-stream format is proposed to take part of our approach and the associated decoding architecture was designed. We analyzed the compression induced by our coding method and proved that compression ratios of at least $2.5\times$ can be achieved on the 20 largest MCNC benchmarks. The introduction of clustering which aggregates multiple routing resources together showed compression ratio up to a factor of $10\times$, at the cost of a more complex decoding step at runtime. Future perspectives on the VBS include extension of the architecture to support commercially available FPGAs as well as the improvement of the associated CAD tool flow to include smarter coding of the VBS to gain in runtime efficiency and in size. The VBS approach can provide increased online relocation capabilities using a decoding algorithm capable of decoding the VBS on-the-fly during the task migration. We applied for a European Patent on this work [73] and the results will be published in 2015 at IEEE/ACM DATE [47].

6.2.2. Power Models of Reconfigurable Architectures

Participants: Robin Bonamy, Daniel Chillet, Olivier Sentieys.

Including a reconfigurable area in complex systems-on-chip is considered as an interesting solution to reduce the area of the global system and to support high performance. But the key challenge in the context of embedded systems is currently the power budget and the designer needs some early estimations of the power consumption of its system. Power estimation for reconfigurable systems is a difficult issue since several parameters need to be taken into account to define an accurate model. In this research, we consider the opportunity of the dynamic reconfiguration for the reduction of power consumption by the management of tasks scheduling and placement. We analyzed the power consumption during the dynamic reconfiguration on a Virtex 5 board. Three models of the partial and dynamic reconfiguration power consumption with different complexity/accuracy tradeoffs are defined. These models are used in design space exploration to evaluate the impact of reconfiguration on energy consumption of a complete system. We propose a methodology for power/energy consumption modeling and estimation in the context of heterogeneous (multi)processor(s) and dynamically reconfigurable hardware systems. We developed an algorithm to explore all task mapping possibilities for a complete application (e.g., for H264 video coding) with the aim to extract one of the best solutions with respect to the designer's requirements. This algorithm is a step ahead for defining on-line power management strategies to decide which task instances must be executed to efficiently manage the available power using dynamic partial reconfiguration [24].

6.2.3. Real-time Spatio-Temporal Task Scheduling on 3D Architecture

Participants: Quang-Hai Khuat, Quang Hoa Le, Emmanuel Casseau, Antoine Courtay, Daniel Chillet.

One of the main advantages offered by a three-dimensional system-on-chip (3D SoC) is the reduction of wire length between different blocks of a system, thus improving circuit performance and alleviating power overheads of on-chip wiring. To fully exploit this advantage, an efficient management referring to allocate temporarily the tasks at different levels of the architecture is greatly important. In the context of 3D SoC, we have developed several spatio-temporal scheduling algorithms for 3D MultiProcessor Reconfigurable System-on-Chip (3DMPRSoC) architectures composed of a multiprocessor layer and an embedded Field Programmable Gate Array (eFPGA) layer with dynamic reconfiguration. These two layers are interconnected vertically by through-silicon vias (TSVs) ensuring tight coupling between software tasks on processors and associated hardware accelerators on the eFPGA. Our algorithms cope with task dependencies and try to allocate communicating tasks close to each other in order to reduce direct communication cost, thus reducing global communication cost. In the 3DMPRSoC context, our algorithms favor direct communications including: i) point-to-point communication between hardware accelerators on the eFPGA, ii) communication between software tasks through the Network-on-Chip of the multiprocessor layer, and iii) communication between software task and accelerator through TSV. When a direct communication between two tasks occurs, the data are stored in a shared memory placed onto the multiprocessor layer.

The algorithm proposed in [50] considers heterogenous reconfigurable architecture and proposes a mathematical formulation for spatio-temporal scheduling of a task graph. The placement consists in finding the best mapping of the application task model onto the reconfigurable region. To improve the performance of our algorithm, we propose to configure the tasks by taking account of their priority. The global objective consists in the reduction of the global execution time. The second algorithm presented in [51] improves the previous one and proposes to exploit the presence of processor in the multiprocessor layer in order to anticipate a software execution of a task when no sufficient area is available. In this case, classical algorithms reject the task, and continue their execution. Our algorithm starts a software execution of the task, but the software execution is a speculative execution. Indeed, if a sufficient area is freed by a hardware task later, in this case our algorithm evaluates if the software execution must continue or if it is better to stop this execution to restart the task in the reconfigurable area. We demonstrated that the execution time of an application can be significantly reduced by applying this software speculation.

In [53], we proposed a heuristic which focus on the online task placement problem on a multi-context, dynamically and partially reconfigurable heterogeneous architecture. Configuration prefetching and anti-fragmentation well known techniques are combined with the place reservation technique that takes into account tasks to be placed in the future (pre-allocated tasks) while fulfilling task execution deadline constraint. Compared to a placement without reservation, our approach improves the number of placed tasks and the resource utilization rate.

6.2.4. Run-time Task Management to Increase Resource Utilisation for Concurrent Critical Tasks in Mixed-Critical Systems

Participant: Angeliki Kritikakou.

When integrating mixed critical systems on a multi/many-core system, one challenge is to ensure predictability for the high criticality tasks and an increased utilization for low criticality tasks. In [52], we proposed a distributed run-time WCET controller to address this problem, when several high criticality tasks with different deadlines, periods and offsets are concurrently executed on a multi core system.

During the system execution, the proposed controller regularly checks locally at each critical task if the interferences due to the low criticality tasks can be tolerated. This is achieved by monitoring the ongoing execution time, dynamically computing the remaining worst case execution time of the critical task when only critical tasks are executed on the system and checking our safety condition. In case that the condition is violated for one critical task, the concurrent execution of the low criticality tasks with the critical one will lead to its deadline miss. Therefore, the local controller decides the suspension of the less critical tasks. However, the local controller is not responsible for the actual suspension of the low criticality tasks. The controller sends a request to a master which has a global view of the system. The master is in charge of collecting the requests of the critical tasks, suspending and restarting the low criticality tasks. When at least one critical task sends the request for suspension of the low criticality tasks, the master suspends them. During execution, the master updates the number of active requests and it restarts the low criticality tasks when all requesters have finished their execution. We have implemented our approach as a software controller on a real multi-core COTS system, the TMS320C6678 chip of Texas Instruments, where we have observed significant gains up to 556% for our case study.

6.2.5. Arithmetic Operators for Cryptography and Fault-Tolerance

Participants: Arnaud Tisserand, Emmanuel Casseau, Nicolas Veyrat-Charvillon, Karim Bigou, Franck Bucheron, Jérémie Métairie, Gabriel Gallin, Huu Van Long Nguyen, Nicolas Estibals.

Arithmetic Operators for Fast and Secure Cryptography.

In the paper [39] presented at ASAP, we describe a new RNS (residue number system) modular multiplication algorithm, for finite field arithmetic over GF(p), based on a reduced number of moduli in base extensions with only $3n/2$ moduli instead of $2n$ for standard ones. Our algorithm reduces both the number of elementary modular multiplications (EMMs) and the number of stored precomputations for large asymmetric cryptographic applications such as elliptic curve cryptography or Diffie-Hellman (DH) cryptosystem. It leads to faster operations and smaller circuits.

The PhD thesis defended by Karim Bigou [16] deals with the RNS representation and the associated arithmetic algorithms for asymmetric cryptography (ECC and RSA). The title of the PhD is "Theoretical Study and Hardware Implementation of Arithmetical Units in Residue Number System (RNS) for Elliptic Curve Cryptography".

Scalar recoding is popular to speed up ECC (elliptic curve cryptography) scalar multiplication: non-adjacent form, double-base number system, multi-base number system (MBNS). Ensuring uniform computation profiles is an efficient protection against some side channel attacks (SCA) in embedded systems. Typical ECC scalar multiplication methods use two point operations (addition and doubling) scheduled according to secret scalar digits. Euclidean addition chains (EAC) offer a natural SCA protection since only one point operation is used. Computing short EACs is considered as a very costly operation and no hardware implementation has been reported yet. We designed an hardware recoding unit for short EACs which works concurrently to scalar multiplication. It has been integrated in an in-house ECC processor on various FPGAs. The implementation results show similar computation times compared to non-protected solutions, and faster ones compared to typical protected solutions (e. g. 18 % speed-up over 192 b Montgomery ladder).

In the paper [40], we introduce a robust asynchronous logic family which does not rely on timing assumptions and/or delay elements and can operate with sub-powered devices. The key element behind our proposal is a simplified completion detection mechanism which makes it substantially more energy effective when compared with other dual-rail approaches. A 32-bit Ripple Carry Adder (RCA) is implemented in 65nm and 45nm CMOS process to evaluate the practicability of our approach. Firstly, the Optimal Energy Point (OEP) of the proposed RCA is investigated by scaling VDD from 0.4V to 0.2V (50mV interval), where the OEP occurs at 0.25V for both technologies. Secondly, while comparing the energy consumption with the corresponding single-rail benchmark at its OEP in 65nm process, 30% (34 fJ for 65nm) and 40% (54fJ for 45nm after scaling) energy savings are achieved respectively. More impressive (10x better) energy efficiency and reasonable performance are obtained over dual-rail counterparts. This work is done in the SPiNaCH project.

ECC Crypto-Processor with Protections Against SCA.

A dedicated processor for elliptic curve cryptography (ECC) is under development. Functional units for arithmetic operations in GF(2^m) and GF(p) finite fields and 160-600-bit operands have been developed for FPGA implementation. Several protection methods against side channel attacks (SCA) have been studied. The use of some number systems, especially very redundant ones, allows one to change the way some computations are performed and then their effects on side channel traces. This work is done in the PAVOIS project.

Arithmetic Operators and Crypto-Processor for HECC.

In the HAH project, we study and prototype efficient arithmetic algorithms for hyperelliptic curve cryptography for hardware implementations (on FPGA circuits). We study new advanced arithmetic algorithms and representations of numbers for efficient and secure implementations of HECC in hardware.

Arithmetic Operators for Fault Tolerance.

In the ARDyT and Reliasic projects, we work on computation algorithms, representations of numbers and hardware implementations of arithmetic operators with integrated fault detection (and/or fault tolerance) capabilities. The target arithmetic operators are: adders, subtracters, multipliers (and variants of multiplications by constants, square, FMA, MAC), division, square-root, approximations of the elementary functions. We study two approaches: residue codes and specific bit-level coding in some redundant number systems for fault detection/tolerance integration at the arithmetic operator/unit level. FPGA prototypes are under development.

Secure Virtualization in Hardware

In the paper [70] presented at SDTA, we deal with secure solutions that can help virtualization and communication which can be implemented on new hybrids (Core + FPGA) development platforms. On one side, these boards are featured with processors that do not have virtualization extensions but are powerfull enough to really support hypervisors and their guests. On the other side some virtualization solutions presently exist for ARM processors but they only refer to TrustZone for their (hardware) security. These hybrid boards can offer us more: we have read some recents and up-to-date specifications made by a consortium to help the implementation of hardware security. In this area, FPGA can help in securing virtualization. But we must notice that, for now, all has been made for Intel/AMD architectures and for a lone operating system. Even so, the whole propositions are too complex to be implemented on embedded systems. So, we will have to use some capabilities in hardware development and make software rearrangements to help us to design a functional solution.

6.3. Compilation and Synthesis for Reconfigurable Platform

6.3.1. Numerical Accuracy Analysis and Optimization

Participants: Olivier Sentieys, Steven Derrien, Romuald Rocher, Pascal Scalart, Tomofumi Yuki, Aymen Chakhari, Gaël Deest.

The problem of accuracy evaluation is one of the most time consuming tasks during the fixed-point refinement process. Analytical techniques based on perturbation theory have been proposed in order to overcome the need for long fixed-point simulation. However, these techniques are not applicable in the presence of certain operations classified as un-smooth operations. In such circumstances, fixed-point simulation should be used. In [33], an algorithm detailing the hybrid technique which makes use of an analytical accuracy evaluation technique used to accelerate fixed-point simulation was proposed. This technique is applicable to signal processing systems with both feed-forward and feedback interconnect topology between its operations. The proposed algorithm makes use of the classification of operators as smooth or un-smooth and uses the analytical SNS model obtained by using our previously published analytical techniques to evaluate the impact of finite precision on smooth operators, while performing simulation of the un-smooth operators during fixed-point simulation. In other words, parts of the system are selectively simulated only when un-smooth errors occur and not otherwise. Thus, the effort for fixed-point simulation is greatly reduced. The acceleration obtained as a result of applications of the proposed technique is consistent with fixed-point simulation, while reducing the time taken for fixed-point simulation by several orders of magnitude. The preprocessing overhead consists of deriving the single-noise-source model, and it is often small in comparison to the time required for fixed-point simulation. The advantage of using the proposed technique is that the user need not spend time on characterizing the nonlinearities associated with un-smooth operations. Several examples from general signal processing, communication, and image processing domains are considered for evaluation of the proposed hybrid technique. The acceleration obtained is quantified as an improvement factor. Very high improvement factors indicate that the hybrid simulation is several orders of magnitude faster than classical fixed-point simulation.

One of the limitation of analytical accuracy technique is that they are based on a Signal Flow Graph Representation of the system to be analyzed. This SFG model is currently built-out of a source program by flattening its whole control-flow (including full loop unrolling) which raises significant accuracy analysis issues. To overcome these limitations, we have proposed [41] to adapt state of the art accuracy analysis techniques to take advantage of compact polyhedral program representations. Combining the two approaches provide a more general and scalable framework which significantly extends the applicability of accuracy models, enabling the analysis of complex image processing kernels operating on multidimensional data-sets.

An analytical approach was studied to determine accuracy of systems including unsmooth operators. An unsmooth operator represents a function which is not derivable in all its definition interval (for example the sign operator). The classical model is no longer valid since these operators introduce errors that do not respect the Widrow assumption (their values are often higher than signal power). So an approach based on the distribution of the signal and the noise was proposed. We focused on recursive structures where an error influences future decision (such as Decision Feedback Equalizer). In that case, numerical analysis method (e.g., Newton Raphson algorithm) can be used. Moreover, an upper bound of the error probability can be analytically determined. We also studied the case of Turbo Coder and Decoder to determine data word-length ensuring sufficient system quality [17].

6.3.2. Reconfigurable Processor Extension Generation

Participants: Christophe Wolinski, François Charot.

Most proposed techniques for automatic instruction sets extension usually dissociate pattern selection and instruction scheduling steps. The effects of the selection on the scheduling subsequently produced by the compiler must be predicted. This approach is suitable for specialized instructions having a one-cycle duration because the prediction will be correct in this case. However, for multi-cycle instructions, a selection that does not take scheduling into account is likely to privilege instructions which will be, *a posteriori*, less interesting than others in particular in the case where they can be executed in parallel with the processor core. The originality of our research work is to carry out specialized instructions selection and scheduling in a single optimization step. This complex problem is modeled and solved using constraint programming techniques. This approach allows the features of the extensible processor to be taken into account with a high degree of flexibility. Different architectures models can be envisioned. This can be an extensible processor tightly coupled to a hardware extension having a minimal number of internal registers used to store intermediate results, or a VLIW-oriented extension made up of several processing units working in parallel and controlled by a specialized instruction. These techniques have been implemented in the Gecos source-to-source framework.

Novel techniques addressing the interactions between code transformation (especially loops) and instruction set extension are under study. The idea is to automatically transform the original loop nests of a program (using the polyhedral model) to select specialized and vector instructions. These new instructions may use local memories located in the hardware extension and used to store intermediates data produced at a given loop iteration. Such transformations lead to patterns whose effect is to significantly reduce the pressure on the memory of the processor.

We also studied a way to identify custom instructions at the application domain level instead of addressing it on a per-application basis. Domain-specific instruction set extension aims at maximizing the usage of a custom instruction across a set of applications belonging to an application domain. The idea is to guarantee that each custom instruction has a high degree of utilization across many applications of a given domain, while still delivering the required performance improvement. The instruction identification problem is here formulated as the maximum common subgraph problem and it is solved by transforming it into a maximum clique problem.

6.3.3. Optimization of Loop Kernels Using Software and Memory Information

Participant: Angeliki Kritikakou.

The compilers optimize the compilation sub-problems one after the other following an order which leads to less efficient solutions because the different sub-problems are independently optimized taking into account only a part of the information available in the algorithms and the architecture. In a paper accepted for publication in Computer Languages, Systems & Structures (COMLAN), Elsevier, we have presented an approach which applies loop transformations in order to increase the performance of loop kernels. The proposed approach focuses on reducing the L1, L2 data cache and main memory accesses and the addressing instructions. Our approach exploits the software information, such as the array subscript equations, and the memory architecture, such as the memory sizes. Then, it applies source-to-source transformations taking as input the C code of the loop kernels and producing a new C code which is compiled by the target compiler. We have applied our approach to five well-known loop kernels for both embedded processors and general purpose processors. From the obtained experimental results we observed speedup gains from 2 up to 18.

6.3.4. Design Tools for Reconfigurable Video Coding

Participants: Emmanuel Casseau, Yaset Oliva Venegas.

In the field of multimedia coding, standardization recommendations are always evolving. To reduce design time taking benefit of available SW and HW designs, Reconfigurable Video Coding (RVC) standard allows defining new codec algorithms. The application is represented by a network of interconnected components (so called actors) defined in a modular library and the behaviour of each actor is described in the specific RVC-CAL language. Dataflow programming, such as RVC applications, express explicit parallelism within an application. However general purpose processors cannot cope with both high performance and low power consumption requirements embedded systems have to face. We have investigated the mapping of RVC applications onto a dedicated multiprocessor platform. Actually, our goal is to propose an automated co-design flow based on the RVC framework. The designer provides the application description in the RVC-CAL language, after which the co-design flow automatically generates a network of processors that can be synthesized on FPGA platforms. Two kinds of platforms can be targeted. The first platform is made of processors based on a low complexity and configurable TTA processor (Very Long Instruction Word -style processor). The architecture model of the platform is composed of processors with their local memories, an interconnection network and shared memories. Both shared and local memories are used to limit the traditional memory bottleneck. Processors are connected together through the shared memories [72] [69] [36]. The second platform more specifically targets the Zynq platform from Xilinx. The processors are MicroBlaze processors. Their local memory is dedicated to instruction code only. A common shared memory is used for the data exchanges between the processors (to store the data that communicate between actors). At present time, the actor mapping is chosen at compile time but we expect dynamic mapping soon. The mapping will be computed at runtime on the ARM processor. The actor's code will be stored in the DDR memory so that it can be easily transferred to the MicroBlaze instruction cache depending on the actor mapping [55] [76] . This work is done in collaboration with IETR and has been implemented in the Orcc open-source compiler (Open RVC-CAL Compiler: http://orcc.sourceforge.net).

6.3.5. A Domain Specific Language for Rapid Prototyping of Software Radio Waveforms

Participants: Matthieu Gautier, Olivier Sentieys, Ganda-Stéphane Ouedraogo.

Software Defined Radio (SDR) is now becoming a ubiquitous concept to describe and implement Physical Layers (PHYs) of wireless systems. Moreover, even though the FPGA (Field Programmable Gate Array) technology is expected to play a key role in SDR, describing a PHY at the Register-Transfer-Level (RTL) requires tremendous efforts. We introduced a novel methodology to rapidly implement PHYs for FPGA-SDR platforms. The work relies upon High-Level Synthesis tools and dataflow modeling to infer an efficient system-level control unit for the application. The proposed software-based over-layer partly handles the complexity of programming an FPGA and integrates reconfigurable features. It consists essentially of a Domain-Specific Language (DSL) [60] that handles the complexity of programming an FPGA and a DSL-Compiler [32] for automation purpose. IEEE 802.11a a and IEEE 802.15.4 transceivers have been designed and explored [45] via this new methodology in order to show the rapid prototyping feature.

6.4. Interaction between Algorithms and Architectures

6.4.1. Cooperative-cum-Constrained Maximum Likelihood Algorithm for UWB-based Localization in Wireless BANs

Participants: Antoine Courtay, Matthieu Gautier, Gia Minh Hoang [Master's Student].

Wireless Body Area Network (BAN) is a mainstream technology for numerous application fields (medicine, security, sport science, etc.) and precise determination of wireless sensors' positions responses to the great needs in many applications. This study leverages Ultra Wide Band (UWB) radio which is an attractive technology to achieve the centimeter-level distance measurements. However, the aggregation of the distance information remains a challenge to achieve an accurate localization in wireless BAN. To this aim, we have proposed a novel Cooperative-cum-Constrained Maximum Likelihood (CCML) localization algorithm. This algorithmic study shows the improvement that could be achieved by combining UWB radio and dedicated algorithms. Future works is to integrate UWB technology in the second version of the Zyggie platform developed in CAIRN.

6.4.2. MIMO Systems and Cooperative Strategies for Low-Energy Wireless Networks

Participants: Olivier Berder, Olivier Sentieys, Baptiste Vrigneau, Viet-Hoa Nguyen.

Since a couple of years, the CAIRNteam has reached a significant expertise in multi-antenna systems, especially in linear precoding. If this technique is traditionally used in a collocated way, it could also be used for wireless sensor networks (WSN) in a distributed manner. We presented a new approach, named distributed max-dmin precoding (DMP). This protocol is based on the deployment of a virtual 2x2 max-dmin precoding over one source, one forwarding relay, both equipped with one antenna and a destination involving two antennas. In this context, two kinds of relaying, amplify and forward or decode and forward protocols, were investigated. The performance evaluation in terms of Bit-Error-Rate (BER) and energy efficiency was compared with non cooperative techniques (SISO, SIMO) and the distributed space time block code (STBC) scheme. Our investigations showed that the DMP takes the advantage in terms of energy efficiency from medium transmission distances.

A receiver initiated cooperative medium access control (RIC-MAC) protocol was also proposed for cooperative communications to reduce the energy consumption of WSN. Considering a real WSN platform, the simulation results show that using the proposed RIC-MAC protocol in cooperative communications provides latency and energy gains as compared to multi-hop communications. Even if the energy gain is shown to be reduced when the network traffic load increases, our protocol still brings an energy gain about 22% at 1 packet/second. Finally, considering the impact of traffic load on energy consumption and latency, RIC-MAC is illustrated to be robust to traffic load variations in terms of latency [66].

6.4.3. Adaptive protocols for Wireless Sensor Networks

Participants: Olivier Berder, Matthieu Gautier, Nhat-Quang Nhan [Master's Student], Van-Thiep Nguyen.

As tiny sensor nodes are equipped with limited battery, the optimization of the power consumption of these devices is extremely vital. In typical WSN platforms, the radio transceiver consumes major proportion of the energy. Major concerns are therefore to decrease the radio activity by designing efficient MAC protocols.

Energy consumption plays an important role in the design of Wireless Body Area Sensor Network (WBASN). Unfortunately, the performance of WBASNs decreases in high interference environments such as the Industrial, Scientific and Medical (ISM) band where wireless spectrums are getting crowded. In this study [59], an energy-efficient Medium Access Control (MAC) protocol named C-RICER (Cognitive-Receiver Initiated CyclEd Receiver) is specifically designed for WBASN to cognitively work in high interference environment. C-RICER protocol adapts both transmission power and channel frequency to reduce the interferences and thus, the energy consumption. The protocol is simulated with the OMNET++ simulator. Simulation results show that, depending on the interference level, C-RICER is able to outperform the traditional RICER protocol in terms of energy consumption, packet delay, and network throughput.

In recent years, many MAC protocols for Wireless Sensor Networks (WSNs) have been proposed and evaluated using Matlab simulator and/or network simulators (OMNeT++, NS2, etc.). However, most of them have a static behavior and few network simulations are available for adaptive protocols. Specially, in OMNeT++/MiXiM, there is few energy-efficient MAC protocol for WSNs (B-MAC and L-MAC) and no adaptive protocol. To this end, the TAD-MAC (Traffic Aware Dynamic MAC) protocol has been simulated in OMNeT++ with the MiXiM framework [57]. The simulation results have been used to compare with B-MAC and L-MAC protocol, showing the gain brought by TAD-MAC.

6.4.4. Energy Harvesting and Power Management

Participants: Olivier Berder, Olivier Sentieys, Arnaud Carer, Trong-Nhan Le.

To design autonomous Wireless Sensor Networks (WSNs) with a theoretical infinite lifetime, energy harvesting (EH) techniques have been recently considered as promising approaches. Ambient sources can provide everlasting additional energy for WSN nodes and exclude their dependence on battery. An efficient energy harvesting system which is compatible with various environmental sources such as light, heat or wind energy was proposed. Our platform takes advantage of double-level capacitors not only to prolong the system lifetime but also to enable robust booting from the exhausting energy of the system. Simulations and experiments showed that it can achieve booting time in order of seconds. Although capacitors have virtual recharge cycles, they suffer from higher leakage compared to rechargeable batteries. Increasing their size can decrease the system performance due to leakage energy. Therefore, an energy neutral design framework providing a methodology to determine the minimum size of the storage devices satisfying Energy Neutral Operation (ENO) and maximizing system Quality of Service (QoS) in EH nodes when using a given energy source was proposed. Experiments validating this framework were performed on a real WSN platform with both photovoltaic cells and thermal generators in an indoor environment [30].

A new PM for EH-WSNs scavenging energy from periodic sources, i.e., ambient energy is not available during the full harvesting cycle, was proposed. Not only respecting the ENO condition, our PM is able to balance the Quality of Service (QoS) during the whole cycle to provide regular data tracking, which is essential for WSN applications like monitoring. Simulations on OMNET++ show that our PM can improve the QoS during the absence of energy by a factor up to 84% compared to state-of-the-art PMs, while guaranteeing the same global QoS [54].

6.4.5. Multimedia Processing

Participant: Pascal Scalart.

Most noise reduction methods for multimedia signals are usually based on the application of a short-time Wiener filter (MMSE) that is generally expressed as a spectral gain depending on the local signal-to-noise ratio (SNR) on each frequency bin. To estimate such filter, several algorithms can be found in the literature but these conventional approaches lead to a biased estimator for the a priori signal-to-noise estimate. To reduce this bias, we have proposed in [26] a new strategy that relies on the introduction of a correction term in the computation of the Wiener filter depending on the current state of both the available a priori and a posteriori SNR estimates. The proposed solution leads to a bias-compensated a priori SNR estimate, and allows to finely estimating the target signal that is very close to the original noise-free reference. Such refinement procedure has been tested under various noisy environments and show the superiority of the proposed strategy compared to competitive algorithms.

Audio classification systems have recently gained interest for the design of various real-world multimedia services such as audio database indexing with musical genre classification, video indexing using the soundtrack or context awareness. A large majority of audio classification systems can be viewed as offline applications in the sense that there is no strong restriction about how the signal to be classified is accessed. In [44], we investigate the case where the classification task is performed in real-time in a low-latency classification framework. We proposed different methodologies for the use of feature integration that are based on three key aspects: the selection of the features which have to be temporally integrated, the choice of the integration techniques, i.e. how the temporal information is extracted, and the size of the integration window. The experiments carried out for the classification task show that these different methodologies have a significant impact on the global performance even with the low-latency constraints. In addition, we investigate the detection of howlings that arise in audio signals in [43]. To do so, the processing algorithm is based on a Support Vector Machine (SVM) model in the decision stage and on the combination of energy-based features and also a new feature related to the frequency stability of a howling component. The proposed method can be used in different situation since its provides good results with a very low false alarm rate for a wide range of experimental conditions.

6.4.6. Non-Intrusive Load Monitoring

Participants: Olivier Sentieys, Baptiste Vrigneau, Xuan Chien Le.

Natural resource preservation has recently become a significant concern and has therefore motivated many research and development efforts for energy consumption management in buildings and homes. Efficiently reducing energy consumption at home, work or in a factory, could be afforded by mixing different technologies to not only reduce the energy consumed by consumers, but also to adapt (manage) the energy consumed to the energy that is produced. SMART 2020 outlined the opportunity to capture savings of both energy and Greenhouse Gas (GHG) emissions in 2020, through a range of actions developed by the Information and Communications Technologies (ICT) sector. Smart Grid, Smart Buildings, and Green ICT have the main impact on energy savings. At the energy production side, the electrical grid infrastructure is comprised of three elements: power generation, transmission, and distribution. Electrical power generation consists mainly of the power plants but also includes more and more renewable sources such as wind power or solar panels on energy farms or locally on top of buildings. The cost of energy storage is very high, and hence the current practice is to match energy consumption closely with energy generation, which is more and more fluctuating: challenges could be seen as being able to use energy when the wind blows or the sun shines, and also to avoid the strong power consumption peaks due to people's life. A typical example at home could be to automatically use the dryer when energy is available and therefore cheap, and is now well defined as Smart Grid technologies. At the energy consumption side, the main objective is of course to reduce energy consumption of the different subsystems. Interior lighting, office equipment, heating, cooling, and ventilation make up of more than 85% of the total electricity use and the reduction effort should therefore be concentrated on these systems. For energy management and reduction in homes or building a key enabler is the use of wireless sensor networks to monitor the environment (temperature, activity of people, power consumption of equipment, light, etc.) and to act on subsystems (decrease room temperature, stop or start an equipment, adjust cooling or ventilation, etc.). This is the emerging field of Smart Building Automation.

The objective of this work is strongly linked to the usage of these WSN nodes in the context of smart monitoring of energy consumption and environment (temperature, activity, light). We will propose new Indirect Power Monitoring techniques which enable to estimate energy consumed in a building or in a home without effectively measuring the power consumed. A typical AC smart meter is costly equipment and we therefore want to propose cheap and non-invasive sensor nodes. As an example, to estimate the power consumed by the TV, it is not necessary to measure precisely the current it consumed, but a simple sensor able to recognize that TV is on or off can do the same job with a far less complexity. Another example is the development and deployment of room occupancy and people activity sensors that can lead to significant reduction of the energy by regulating HVAC (Heating, Ventilation and Air-Conditioning) or by switching lights and office equipment. The wireless transmission is the main reason of consuming energy and the new algorithms will propose to make the sensors to cooperate inside a low-distance cluster (an office for example). The algorithms will decide the best strategy and the best information to send back in order to offer the best trade-off between Performance/Complexity/Consumption. This work is closely links to power management techniques and energy harvesting (in-door light, heat, vibration). A power manager embedded in energy harvesting WSN nodes adapts the power consumption and computation loads according to the harvested energy to obtain a theoretically infinite lifetime. The main advantage of using energy harvesting (EH) in the context of building and home monitoring is to avoid battery replacement and therefore to reduce installation and maintenance costs of the system.

CAMUS Team

6. New Results

6.1. Highlights of the Year

One of Philippe Clauss' early papers on Ehrhart polynomials has been celebrated, 18 years later, in a selection of papers for the International Conference on Supercomputing (ICS) 25th anniversary retrospective [13]. 35 papers have been selected among roughly 1800 papers published between 1987 and 2011. The paper is:

"Counting Solutions to Linear and Nonlinear Constraints Through Ehrhart Polynomials: Applications to Analyze and Transform Scientific Programs", by Philippe Clauss, ICS'96, which introduced Ehrhart polynomials in the field of program analysis and optimization.

Philippe Clauss wrote an additional retrospective [12] related to this research which complements the paper in the ICS special issue.

6.2. APOLLO (Automatic speculative POLyhedral Loop Optimizer)

The goal of the APOLLO project is to provide a set of annotations (pragmas) that the user can insert in the source code to perform advanced analyses and optimizations, for example dynamic speculative parallelization. It is based on the prototype named VMAD which was developed previously by the team between 2009 and 2012. Alexandra Jimborean defended her PhD thesis on this topic in 2012 [30].

APOLLO includes a modified LLVM compiler and a runtime system. The program binary files are first generated by our compiler to include necessary data, instrumentation instructions, parallel code skeletons, and callbacks to the runtime system which is implemented as a dynamic library. External modules associated to specific analyses and transformations are dynamically loaded when required at runtime.

APOLLO uses sampling, multi-versioning and code skeletons to limit the runtime overhead (profiling, analysis, and code generation). At runtime, targeted codes are launched by successive chunks that can be either original, instrumented or optimized/parallelized versions. These latter versions are generated on-the-fly through fast instantiation of the code skeletons. After each chunk execution, decisions can be taken relatively to the current optimization strategy. APOLLO is handling advanced memory access profiling through linear interpolation of the addresses, dynamic dependence analysis, version selection and speculative polyhedral parallelization [9].

Several extensions and improvements have been implemented inside Apollo in 2014:

- the scheduler of the polyhedral compiler Pluto has been integrated inside the framework. Thus, the runtime decision regarding what optimizing and parallelizing transformation is now entirely depending on Pluto, whose input is generated by the instrumentation and interpolation phase of Apollo [20].
- the static compilation phase of Apollo has been significantly enforced. Linear dependencies between values of scalars and memory addresses are identified in order to alleviate the cost of the instrumented code version. Additionally, memory reference functions that can be disambiguated at compile-time are now fully handled. These improvements are using analysis passes of the LLVM compiler, as well as passes that were specifically developed.
- Apollo is now using the LLVM JIT compiler to further optimize the instantiated code skeletons. Previously, code skeletons were generated as binary executable at compile-time with global variables instantiated at runtime. This approach yielded sub-optimal code including unnecessary or invariant computations. Code skeletons are now kept in LLVM intermediate form until being instantiated and compiled at runtime using the LLVM JIT compiler, thus resulting in faster optimized codes.

- Other memory behavior modeling approaches are now being studied and implemented, in order to allow Apollo handling codes that do not have a completely linear behavior. Three main cases are addressed:
  - quasi-linear behavior in which memory accesses which do not fit the linear prediction are checked on-the-fly, i.e., if these delinquent accesses do not invalidate the current parallel schedule.
  - linear regression behavior in which memory accesses are staying inside a “tube” bordered by linear functions.
  - behavior in which memory accesses are staying inside disjointed address ranges.

6.3. The XFOR programming structure

We have proposed a new programming control structure called “xfor” or “multifor”, providing users a way to schedule explicitly the statements of a loop nest, and take advantage of optimization and parallelization opportunities that are not easily attainable using the standard programming structures. This work is the PhD work of Imen Fassi, who started her work in 2013 and who is co-advised by Yosr Slama, Assistant Professor at the University El Manar in Tunis, Tunisia, and Philippe Clauss.

Data locality optimization is a well-known goal when handling programs that must run as fast as possible or use a minimum amount of energy. However, usual techniques never address the significant impact of numerous stalled processor cycles that may occur when consecutive load and store instructions are accessing the same memory location. In [15], we show that two versions of the same program may exhibit similar memory performance, while performing very differently regarding their execution times because of the stalled processor cycles generated by many pipeline hazards. The xfor structure enables the explicit control of the way data locality is optimized in a program and thus, to control the amount of stalled processor cycles. In [15], we also show the benefits of xfor regarding execution time and energy saving.

While many advanced and fully automatic program analysis and optimization techniques have been developed thanks to the accuracy and expressiveness of the polyhedral model, these techniques may fail in producing efficient codes in some circumstances. The xfor structure eases the manual application of optimizing transformations on loop nests for expert programmers and allows to generate executable codes that may be significantly faster than those generated automatically using well-established polyhedral strategies. we highlight five main gaps regarding these strategies and discuss some ideas on how to bridge them in [14].

6.4. CPU+GPU adaptive computation

We aim to automatically use CPU and GPU to jointly execute a parallel code. To ensure load balance between different PUs, thus to preserve performance, it is necessary to consider the underlying hardware and the program parameters. Compiler optimizations, execution context, hardware availability and specification make it difficult to determine execution times statically. To overcome this hurdle we rely on a portable and automatic method for predicting execution times of statically generated codes on multicore CPUs and on CUDA GPUs. This approach relies on three stages: automatic code generation, offline profiling of the target code and online prediction.

This is mainly the work of PhD student Jean-François Dollinger, advised by Vincent Loechner since 2011. Preliminary results, a "fastest-wins" algorithm between a multicore CPU and the best predicted GPU code version, was published in 2013 in ICPP. Our latest advances, load balancing code between multiple cores CPUs and multiple GPUs will be presented at the IMPACT 2015 workshop [25] in conjunction with the HiPEAC conference. We are currently preparing an extended journal paper to present this work, and Jean-François Dollinger will defend his PhD in 2015.

6.5. Minimizing the synchronization overhead of X10 programs

The CAMUS team has for long focused on compiling, optimizing, and parallelizing *sequential* programs. The project described in this section is somewhat unusual in this context, in that it targets programs written in an explicitly parallel language, and applies polyhedral modeling techniques to reschedule computations, effectively introducing parallel-to-parallel program transformations. This work has been done in collaboration with the Inria COMPSYS team at ENS Lyon, and first results were presented at the *Compiler Construction* conference (CC'14) in April 2014.

The need to leverage the computing power of multi-core processors (and distributed computers) has lead to the design of explicitly parallel programming languages. Such languages often employ a fork/join model, and include syntax to launch and synchronize tasks (also called activities) with well-defined semantics. This brings parallel constructions under the control of the compiler, and introduces new optimization opportunities. Our work has focused on the various synchronization primitives available to the programmer, and more specifically on how one type of synchronization can be replaced with another for specific classes of programs, the goal being to minimize the synchronization overhead. We have demonstrated significant speedups on programs written using the X10 programming language, and have obtained similar results on equivalent Habanero-Java programs.

More specifically, our work focused on synchronization primitives of X10. The X10 language basically has two activity synchronization primitives: one is the explicit use of “clocks” (synchronization barriers) during activity execution, the other is the implicit use of activity containers that synchronize only on the end of activities. Under reasonable conditions on the patterns of activity creation and control, we showed that long-running activities using clocks can be replaced by short-lived activities synchronized only on the end of their containers, and that this transformation provides a significant gain at run time.

We have studied the converse transformation, i.e. starting with an unclocked X10 program, obtaining a system of sequential threads executing in parallel and synchronizing with clocks. This transformation is interesting since it yields to further optimization opportunities. We have elaborated a system of rules to execute the transformation. Applying these rules to "regular" programs gives good results, but fails on some paradigmatic X10 codes. For irregular programs, some parallelism may be lost. We now are investigating a new set of rules to give a correct result for arbitrary X10 programs. A main difficulty is bringing the proof that the set of upgraded rules will give a correct result.

This work has been done in collaboration with Paul Feautrier, member of the COMPSYS Inria team, in ENS Lyon. The CAMUS team has invited Paul Feautrier one more time for one week in June 2014 in Strasbourg.

6.6. Hardware/Software helper thread prefetching

Heterogeneous Many Cores (HMC) architectures that mix many simple/small cores with a few complex/large cores are emerging as a design alternative that can provide both fast sequential performance for single threaded workloads and power-efficient execution for through-put oriented parallel workloads. The availability of many small cores in a HMC presents an opportunity to utilize them as low-power helper cores to accelerate memory-intensive sequential programs mapped to a large core. However, the latency overhead of accessing small cores in a loosely coupled system limits their utility as helper cores. Also, it is not clear if small cores can execute helper threads sufficiently in advance to benefit applications running on a larger, much powerful, core.

In this project, we designed a hardware/software framework called core-tethering to support efficient helper threading on heterogeneous manycores. Core-tethering provides a co-processor like interface to the small cores that (a) enables a large core to directly initiate and control helper execution on the helper core and (b) allows efficient transfer of execution context between the cores, thereby reducing the performance overhead of accessing small cores for helper execution. Our evaluation on a set of memory intensive programs chosen from the standard benchmark suites shows that helper threads using moderately sized small cores can significantly accelerate a larger core compared to using a hardware prefetcher alone. We find that a small core provides a good trade-off against using an equivalent large core to run helper threads in a HMC. Additionally, helper prefetching on small cores when used along with hardware prefetching, can provide an alternate design point to growing instruction window size for achieving higher sequential performance on memory intensive applications.

This work is a collaboration between the ALF team in Rennes and CAMUS in Strasbourg. Our contribution is mainly on the generation of helper thread code (as a followup to our work on program skeletonization). The result of the work has been published in October 2014 in the Proceedings of the SBAC-PAD conference [17].

6.7. Loop-based Modeling of Parallel Communication Traces

Parallel communication traces are traces of the various actions performed by parallel programs (typically written using MPI or some such library). The traces usually contain actions like message sending and receiving, and entering and exiting collective operations. The goal of this project is to build a model of the parallel program from the traces of the various processes that form the program. Consolidating on our previous work on sequential traces, we have developed an algorithm that takes the traces of the individual processes and merges them into a global model.

The main characteristics of our algorithm is that the result takes the form of loops enclosing various parallel constructs and communication actions. The driving goal of this work is to use the model for various analyzes, mainly to draw qualitative conclusions on the program (like the affinity of the various processes involved), but also to extract quantitative information (like communication matrices). A long term goal is to use the parallel loops to suggest program optimizations.

As of today, our algorithm has been evaluated on several applications. The most obvious is trace compression, with spectacular results because of the underlying loop-nest model (as was already the case for our sequential trace analysis algorithm). Another application is replay, where the program's (actual, i.e., traced) behavior can be simulated on a different parallel architecture. The last application is to build a lightweight model from a subset of trace data, and use the model to index into potentially massive quantitative data associated to the various events.

It turns out that it is difficult to publish such algorithms without evaluating them in "realistic" settings, on applications running on massively parallel hardware, something we don't have easy access to. Also, there are currently a few algorithms that provide similar solutions to practitioners, in a way that we think are fundamentally inferior to our proposition but that seem to be good enough for their current use. Waiting for better opportunities to illustrate the power of our method, we have published a research report summarizing our work [26].

6.8. Switcheable scheduling

Parallel applications used to be executed alone until their termination on partitions of supercomputers. The recent shift to multicore architectures for desktop and embedded systems is raising the problem of the coexistence of several parallel programs. Operating systems already take into account the *affinity* mechanism to ensure a thread will run only onto a subset of available processors (e.g., to reuse data remaining in the cache since its previous execution). But this is not enough, as demonstrated by the large performance gaps between executions of a given parallel program on desktop computers running several processes. To support many parallel applications, advances must be made on the system side (scheduling policies, runtimes, memory management...). However, automatic optimization and parallelization can play a significant role by generating programs with dynamic-auto-tuning capabilities to adapt themselves to the complete execution context, including the system load.

Our approach is to design at compile-time programs that can adapt at run-time to the execution context. The originality of our solution is to rely on *switcheable scheduling*, a selected set of program restructuring which allows to swap between program versions at some meeting points without backtracking. A first step selects pertinent versions according to their performance behavior on some execution contexts. The second step builds the auto-adaptive program with the various versions. Then at runtime the program selects the best version by a low overhead sampling and profiling of the versions, ensuring every computation is useful.

This is an ongoing work with the PhD student Lénaïc Bagnères (POSTALE Team at Inria Saclay-Île-de-France, co-advised by Christine Eisenbeis and Cédric Bastoul). The first results have been presented in 2014 at the Euro-Par International Conference [11].

6.9. Interactive Code Restructuring

This work falls within the exploration and development of semi-automatic programs optimization techniques. It consists in designing and evaluating new visualization and interaction techniques for code restructuring, by defining and taking advantage of visual representations of the underlying mathematical model. The main goal is to assist programmers during program optimization tasks in a safe and efficient way, even if they neither have expertise into code restructuring nor knowledge of the underlying theories. This project is an important step for the efficient use and wider acceptance of semi-automatic optimization techniques, which are still tedious to use and incomprehensible for most programmers. More generally, this research is also investigating new presentation and manipulation techniques for code, algorithms and programs, which could lead to many practical applications: collaboration, tracking and verification of changes, visual search in large amount of code, teaching, etc.

This is a rather new research direction which strengthen CAMUS's static parallelization and optimization issue. It has been initiated at Paris-Sud University as a collaboration between Compilation, represented by Cédric Bastoul before he joined CAMUS, and Human-Machine Interaction, represented by Stéphane Huot from the IN-SITU Team at Inria Saclay-Île-de-France. This work is essentially the PhD topic of Alexander Zinenko (IN-SITU Team at Inria Saclay-Île-de-France, co-advised by Stéphane Huot and Cédric Bastoul, CORDI Grant) which started in 2013. The first results have been presented in 2014 to the IEEE VL/HCC Conference [22]. Moreover, another paper on the topic has been accepted to the International IMPACT 2015 Workshop to be held in conjunction with the HiPEAC International Conference.

CARAMEL Project-Team

6. New Results

6.1. Highlights of the Year

Razvan Barbulescu, ex-PhD student in the team, has received the award “Prix Le Monde de la recherche universitaire”, as one of the top-5 PhD thesis in exact science in 2014.

Emmanuel Thomé has received the “Prix Régional du Chercheur” of the Région Lorraine.

Emmanuel Thomé has received the “Prix de l’Association des Amis de l’Université de Lorraine”.
BEST PAPER AWARD :
[17] **Eurocrypt 2014**. R. BARBULESCU, P. GAUDRY, A. JOUX, E. THOMÉ.

6.2. Discrete logarithm computation in a prime finite field of 180 decimal digits

Participants: Cyril Bouvier, Pierrick Gaudry, Hamza Jeljeli, Emmanuel Thomé [contact].

In the context of the CATREL ANR project, we performed a new computation of a discrete logarithm modulo a 180 digit (596-bit) prime using the number field sieve algorithm. Previous records were 135-digit (448 bits, done in 2006) and 160-digit (530-bit, done in 2007) primes. This is, to date, the largest computation in a prime field. In total, this took the equivalent of 130 years on one CPU core.

6.3. Discrete logarithm in finite fields of small extension degree

Participant: Pierrick Gaudry [contact].

Together with Razvan Barbulescu (CNRS, IMJ-PRG), Aurore Guillevic and François Morain (GRACE project-team), we investigated the discrete logarithm problem in the case of finite fields of the form $\mathbb{F}_{p^n}$, where $n > 1$ is a small integer. We proposed in a preprint — a part of which was accepted to Eurocrypt 2015 — various theoretical and practical improvements [25]:

- new methods for selecting polynomials,
- better (heuristic) asymptotic complexity in the case where $n \approx \log p$, and
- use of algebraic number theory to show that in some cases we can skip the Schirokauer maps.

We have adapted CADO-NFS in order to perform a record computation in a field of the form $\mathbb{F}_{p^2}$, where p^2 has 180 digits. To our knowledge, this is the first time that the number field sieve algorithm is used in practice for record-size computations in this type of fields.

6.4. Igusa class polynomials computation for class number 20,016

Participant: Emmanuel Thomé [contact].

In collaboration with the LFANT project-team, Emmanuel Thomé and Andreas Enge completed the computation of Igusa class polynomials for a quartic CM field whose Igusa class number is 20,016. That is more than 20 times more than the previous state of the art. This has been made possible with the CMH software, which corresponds to the article [10].

6.5. Isogeny graphs for curves with maximal real multiplication

Participant: Emmanuel Thomé [contact].

Emmanuel Thomé and Sorina Ionica (currently with the LFANT project-team) worked on a new algorithm for computing isogeny graphs for Jacobians of curves having the special property that the intersection of their endomorphism ring with its real subfield is maximal. The resulting algorithm is the first depth-first algorithm for this task. This work has been submitted [29].

6.6. Polynomial selection for the Number Field Sieve

Participants: Cyril Bouvier, Nicholas Coxon, Alexander Kruppa, Paul Zimmermann [contact].

A new polynomial selection algorithm for GNFS (General Number Field Sieve) has been described in a preprint [24] and implemented in CADO-NFS. We demonstrate the efficiency of this algorithm by exhibiting a better polynomial than the one used for the factorization of RSA-768, and a polynomial for RSA-1024 that outperforms the best published one.

Montgomery's method of polynomial selection for GNFS has been analysed in a preprint [27]. Criteria for the selection of good parameters for Montgomery's method are given, and the existence of the modular geometric progressions used in the method is considered.

6.7. Beyond double precision

Participant: Paul Zimmermann [contact].

A project entitled “Beyond Double Precision” (BeDoP) has been submitted to the European Research Council (ERC) for funding (advanced grant category). The BeDoP project will (i) demonstrate the limits of double precision on large-scale applications, (ii) make multiple-precision tools easier to use in modern computer languages, and (iii) improve the efficiency and robustness of those tools, in particular by using formal proof techniques. Our dream with the BeDoP project is that scientific computations on exascale computers will no longer give very fast and very wrong results, but instead give very fast and very accurate results.

6.8. Gröbner bases for sparse algebraic systems

Participant: Pierre-Jean Spaenlehauer [contact].

In collaboration with Jean-Charles Faugère and Jules Svartz (POLSYS project-team), new Gröbner bases algorithms have been proposed in [20] to solve efficiently sparse systems of multivariate polynomial equations. Moreover, new complexity bounds have also been proved; they extend in a unified way previous bounds known for solving multi-homogeneous systems with Gröbner bases. For such systems, a proof-of-concept prototype implementation of these algorithms achieves large speed-ups compared to state-of-the-art optimized Gröbner bases algorithms.

6.9. Faster index calculus in algebraic curves

Participant: Maike Massierer [contact].

A possible application of the new ideas speeding up the function field sieve algorithm to index calculus in Jacobians of algebraic curves of large genus has been studied in [30]. Based on a number of practical experiments as well as theoretical considerations, a conjecture has been formulated. It implies that the new ideas only apply to curves which are not interesting in the context of the discrete logarithm problem.

6.10. FFS factory

Participant: Jérémie Detrey [contact].

An extension of Coppersmith's “factorization factory” and Barbulescu's “discrete logarithm factory” to the Function Field Sieve was proposed, dubbed the “FFS factory” [28]. The idea is to batch discrete logarithm computations in finite fields of different extension degrees, sharing the sieving step on the algebraic side between all these finite fields. A careful analysis proved that this approach can be used to lower the overall asymptotic complexity. This was also illustrated with a practical experiment in which the discrete logarithm problem was solved for all of the 50 binary fields of the form $\mathbb{F}_{2^n}$ with n odd ranging from 601 to 699.

CARMEN Team

6. New Results

6.1. Highlights of the Year

- New associated team EPICARD (principal investigator N. zemzemi, Y. Coudière and J. Henry). The aim of of this associated team for the first year is to overcome the technical difficulties that we pointed out during the year 2014 in inverse problem for the heart.
- June 2014: Based on a peer-reviewed proposal, the Grand équipement national de calcul intensif (GENCI) has attributed us 3 million core-hours on the national high-performance computing system Turing, to be used in the year 2014.
- December 2014: Based on a peer-reviewed proposal, the Grand équipement national de calcul intensif (GENCI) has attributed us 3.5 million core-hours on the national high-performance computing machines Turing, Curie, and Occigen, to be used in the year 2015.
- LIRYC will fund a 2-year postdoctoral position on simulation of Brugada syndrome, a rare ECG anomaly predictive of sudden cardiac death in young, apparently healthy subjects. This work will be performed in tight collaboration with clinicians at the Haut-Lévèque hospital

6.2. Inverse problem

We tested our method using synthetic data generated with a highly realistic forward model. Propagating action potentials were generated using a monodomain reaction-diffusion model with a Ten Tusscher 2006 membrane model. An anisotropic human heart model at 0.2-mm resolution was used for this purpose. Torso potentials were then computed from the simulated transmembrane currents using a finite-difference torso model at 1-mm resolution with intracavitary blood, anisotropic myocardium, lungs, and an anisotropic skeletal muscle layer. We simulated 20 cases: 5 single stimuli, 1 dual stimulus and 14 re-entry simulations. From the simulated torso potentials a 200-channel body surface map recording was extracted and used to test the inverse methods. Inverse solutions in terms of epicardial potentials were computed both with MFS and with our .new optimal control approach. With our algorithms, we were able to construct the electrical potential on the heart surface with a very good accuracy in terms of correlation coefficient. Thus, we could accurately reconstruct the activation pattern.

6.3. CEPS

- Integration of some ionic models into CEPS (N. Zemzemi and F. Caro). Those developments will allow us now to use CEPS for the benchmark named *Second N-version Cardiac Electrophysiology Benchmark Specification actual developments*, see (benchmarck) for more details.
- Development of usefull tools for the code (compilation in order to check the depandancies, validation and coverage of the code). 16 test cases are now implementated in CEPS. Those test cases concern unit test case as test for small resolution of linear system (for the FE P1 implemented into the code) and validatation test case as the heat PDE problem. Tests verify also the parallel implementation.
- At this time, the development of the bi-domain model in CEPS is in progress in CEPS with N. Zemzemi.
- First integration of the new model of S. Labarthe initiated during his PhD with L. Colin. This task needs improvment for validation in terms of development.

6.4. Numerical Scheme

Y. Coudière, C. Pierre and R. Turpault wrote some new high order FV schemes. The goal of this study is a future implementation in CEPS.

6.5. Mathematical Model

M. Potse, P.E. Becue and F. Caro wrote a new model for numerical simulations for cardiac electrophysiology at the microscopic scale. We interfer with the LIRYC in order to describe, as much as possible, the interactions between the extra-cellular medium and the intra-cellular medium.

CARTE Project-Team

6. New Results

6.1. Highlights of the Year

Our team made remarkable progress into the understanding of complexity of higher-order functionals. While a robust class of computable functionals exists at any finite type built from $\mathbb{N}$ and $\rightarrow$ (the Kleene-Kreisel functionals), no satisfying complexity classes had been defined so far, except the class BFF of Basic Feasible Functionals. However that class is not a complexity class in the usual sense and does not offer the possibility to define space complexity or non-deterministic time complexity. In his PhD Hugo Férée has developed a non-trivial notion of size for higher-order functionals using game semantics and he has defined a notion of polynomial-time computable functional including BFF but behaving more satisfactorily in several ways. A paper in preparation will gather these results.

6.2. Malware Detection and Program Analysis

- **Complexity Information Flow in a Multi-threaded Imperative Language.** Program resource analysis using tiering based type system has been extended to analyze the time consumed by multi-threaded imperative programs with a shared global memory, which delineates a class of safe multi-threaded programs. In this work presented at TAMC'14 (Theory and Applications of Models of Computation) [22] Jean-Yves Marion and Romain Péchoux have demonstrated that a safe multi-threaded program runs in polynomial time if (i) it is strongly terminating w.r.t. a non-deterministic scheduling policy or (ii) it terminates w.r.t. a deterministic and quiet scheduling policy. As a consequence, we obtain a characterization of the set of polynomial time functions. As far as we know,this is the first characterization by a type system of polynomial time multi-threaded programs
- **A Categorical Treatment of Malicious Behavioral Obfuscation.** In this work presented at TAMC'14 (Theory and Applications of Models of Computation) [23] Romain Péchoux and Thanh Dinh Ta consider malicious behavioral obfuscation through the use of a new abstract model for process and kernel interactions based on monoidal categories. In this model, program observations are considered to be finite lists of system call invocations. In a first step, the authors have shown how malicious behaviors can be obfuscated by simulating the observations of benign programs. In a second step, they have shown how to generate such malicious behaviors through a technique called path replaying and they have extended the class of captured malwares by using some algorithmic transformations on morphisms graphical representation.
- **Malware Message Classification by Dynamic Analysis.** Guillaume Bonfante, Jean-Yves Marion and Thanh Dinh Ta presented to FPS in 2014 a new approach in malware retro-engineering. Usually, either communications, or code is analyzed. Here, the authors take a hybrid perspective. They showed how malware communication can be seen under a language perspective. They tested their idea on real malware and, for instance, showed that the botnet Zeus uses FTP as an underlying network support.
- **Supertagging with Constraints.** The parsing in Natural Language Processing is usually done by statistical analysis. Formal approaches are much more challenging, usually involving hard problems. Guillaume Bonfante, Bruno Guillaume, Mathieu Morey, and Guy Perrier [24] propose a new stream algorithm which discriminates tags in sentences.

6.3. Computability and Complexity

- **Genericity of semi-computable objects.** One of the main goals of computability theory is to understand and classify the algorithmic content of infinite objects, which can be expressed as the difficulty of computing them or as their ability to help solving problems. In establishing this classification one is often led to separate classes of algorithmic complexity and the construction of counter-examples is usually a hard task that requires the use of advanced technics (among which the so-called priority method with finite injury). The difficulty in such a construction is that the constructed object should satisfy two types of requirements going in opposite directions: it should lack algorithmic content but at the same time should be constructible in some way. In other words, these objects live somewhere between *generic* objects (objects with no structure) and *computable* objects (the most constructible objects). While computability theory provides formal notions of genericity, these ones are always incompatible with computability.

  We introduce a new notion of genericity which has two advantages: it is close to plain genericity, and we prove that it is compatible with semi-computability (for a property, being semi-decidable is a semi-computability notion while being decidable is a plain computability notion). The latter result has important consequences: many ad hoc existing constructions are subsumed by this result and then unified, new results can be obtained whenever the new notion of genericity captures the sought properties, and the result clarifies the role of topology in computability theory.

  This work is the sequel of the STACS 2013 paper [19] and is currently submitted [26].

- **Analytical properties of resource-bounded real functionals.** In [14] Hugo Férée, Walid Gomaa and Mathieu Hoyrup extend the results of [52] to non-deterministic complexity. More precisely, we introduce the analytical concepts of essential point and sufficient set for norms over continuous functions and use them to characterize the class of norms that are computable in non-deterministic polynomial time.

- **Call-by-value, call-by-name and the vectorial behaviour of the algebraic λ-calculus.** In this article published in LMCS (Logical Methods in Computer Science) [12], Ali Assaf, Alejandro Díaz-Caro, Simon Perdrix, Christine Tasson and Benoît Valiron examine the relationship between the algebraic lambda-calculus, a fragment of the differential lambda-calculus and the linear-algebraic lambda-calculus, a candidate lambda-calculus for quantum computation. Both calculi are algebraic: each one is equipped with an additive and a scalar-multiplicative structure, and their set of terms is closed under linear combinations. However, the two languages were built using different approaches: the former is a call-by-name language whereas the latter is call-by-value; the former considers algebraic equalities whereas the latter approaches them through rewrite rules. In this paper, they analyse how these different approaches relate to one another. To this end, four canonical languages based on each of the possible choices are proposed: call-by-name versus call-by-value, algebraic equality versus algebraic rewriting. The various languages are simulating each other. Due to subtle interaction between beta-reduction and algebraic rewriting, to make the languages consistent some additional hypotheses such as confluence or normalisation might be required.

- **Real or Natural numbers interpretations and their effect on complexity.** Guillaume Bonfante, Florian Deloup and Antoine Henrot [13] have shown how deep results in algebraic geometry may be read in a complexity perspective. They show that real numbers though they are not well founded can be used as natural numbers are for program interpretations. The argument is based on Positivstellensatz, a major result proved by Stengle.

- **Information carried by programs about the objects they compute.** In computability theory and computable analysis, finite programs can compute infinite objects. Presenting a computable object via any program for it, provides at least as much information as presenting the object itself, written on an infinite tape. What additional information do programs provide? We characterize this additional information to be any upper bound on the Kolmogorov complexity of the object, i.e., it gives an upper bound on size of a shortest program computing the object.

  This problem can be formalized using the two classical models of computation of Markov-computability [61] and Type-2 computability [74], which are the most famous and studied ways of computing with infinite objects. Many celebrated results comparing these models have been developed in the 50's (theorems by Rice, Rice-Shapiro, Kreisel-Lacombe-Schoenfiled/Ceitin, Friedberg) but a complete understanding of their precise relationship has never been obtained. Our results fill this void, identifying the exact relationship between the two models. In particular this relationship enables us to obtain several results characterizing the computational and topological structure of Markov-semidecidable properties.

This work, made in collaboration with Cristóbal Rojas (Santiago) during his visit as an Inria "Chercheur Invité", has been accepted in STACS 2015 [20].

- **Causal Graph Dynamics.** Causal Graph Dynamics extend Cellular Automata to arbitrary, bounded-degree, time-varying graphs. The whole graph evolves in discrete time steps, and this global evolution is required to have a number of physics-like symmetries: shift-invariance (it acts everywhere the same) and causality (information has a bounded speed of propagation). Pablo Arrighi, Emmanuel Jeandel, Simon Martiel (I3S, Univ. Nice-Sophia Antipolis), and Simon Perdrix are investigating the properties of this model. In particular a work on the reversibility of causal graph dynamics has just been submitted in January 2015.
- **The Parameterized Complexity of Domination-type Problems and Application to Linear Codes.** In this article presented at TAMC'14 (Theory and Applications of Models of Computation) [17], David Cattanéo and Simon Perdrix study the parameterized complexity of domination-type problems. (σ, ρ)-domination is a general and unifying framework introduced by Telle: given $\sigma, \rho \subseteq \mathbb{N}$, a set D of vertices of a graph G is (σ, ρ)-dominating if for any $v \in D$, $|N(v) \cap D| \in \sigma$ and for any $v \notin D$, $|N(v) \cap D| \in \rho$. The main result is that for any σ and ρ recursive sets, deciding whether there exists a (σ, ρ)-dominating set of size k, or of size at most k, are both in W[2]. This general statement is optimal in the sense that several particular instances of (σ, ρ)-domination are W[2]-complete (e.g., DOMINATING SET). This result is also extended to a class of domination-type problems which do not fall into the (σ, ρ)-domination framework, including CONNECTED DOMINATING SET and the problem of the minimal distance of a linear code over a finite field.

  To prove the W[2]-membership of the domination-type problems the authors extend the Turing-way to parameterized complexity by introducing a new kind of non-deterministic Turing machine with the ability to perform 'blind' transitions, i.e., transitions which do not depend on the content of the tapes.
- **Quantum Circuits for the Unitary Permutation Problem**. In this paper [18] presented at DCM'14 (New Development in Computational models) and at the Workshop on Quantum Metrology, Interaction, and Causal Structure 2014 (invited talk), Stefano Facchni and Simon Perdrix consider the *Unitary Permutation* problem which consists, given n quantum gates $U_1, ..., U_n$ and a permutation σ of $\{1, ..., n\}$, in applying the quantum gates in the order specified by σ, i.e., in performing $U_{\sigma(n)} \circ ... \circ U_{\sigma(1)}$.

  This problem has been introduced and investigated in [40] where two models of computations are considered. The first is the (standard) model of query complexity: the complexity measure is the number of calls to any of the quantum gates U_i in a quantum circuit which solves the problem. The second model is roughly speaking a model for higher order quantum computation, where quantum gates can be treated as objects of second order. In both model the existing bounds are improved, in particular the upper and lower bounds for the standard quantum circuit model are established by pointing out connections with the *permutation as substring* problem introduced by Karp.

CASCADE Project-Team

5. New Results

5.1. Results

All the results of the team have been published in journals or conferences (see the list of publications). They are all related with the research program (see before) and the research projects (see after):

- More efficient constructions with lattices
- New constructions from pairings
- Delegation of computations
- Analysis of pseudo-random generators
- Advanced primitives for the privacy in the cloud
- Cryptanalysis of symmetric primitives
- New leakage-resilient primitives
- Stronger security with related-key security

CASSIS Project-Team

6. New Results

6.1. Highlights of the Year

Véronique Cortier was one of the two FLoC plenary speakers during the Vienna Summer of Logic [31].

Steve Kremer and Robert Künnemann got a paper accepted at the 35th IEEE symposium on Security and Privacy [45].

The ANR project SEQUOIA has been accepted.

BEST PAPERS AWARDS :

[43] **Software Security and Reliability (SERE)**. E. FOURNERET, J. CANTENOT, F. BOUQUET, B. LEGEARD, J. BOTELLA.

[47] **The 7th International Symposium on Foundations & Practice of Security FPS'2014**. H. H. NGUYEN, A. IMINE, M. RUSINOWITCH.

6.2. Automated Deduction

We develop general techniques which allow us to re-use available tools in order to build a new generation of solvers offering a good trade-off between expressiveness, flexibility, and scalability. We focus on the careful integration of combination techniques and rewriting techniques to design decision procedures for a wide range of verification problems.

6.2.1. Combination of Satisfiability Procedures

Participant: Christophe Ringeissen.

A satisfiability problem is often expressed in a combination of theories, and a natural approach consists in solving the problem by combining the satisfiability procedures available for the component theories. This is the purpose of the combination method introduced by Nelson and Oppen. However, in its initial presentation, the Nelson-Oppen combination method requires the theories to be signature-disjoint and stably infinite (to guarantee the existence of an infinite model). The design of a combination method for non-disjoint unions of theories is clearly a hard task but it is worth exploring simple non-disjoint combinations that appear frequently in verification. An example is the case of shared sets, where sets are represented by unary predicates. Another example is the case of bridging functions between data structures and a target theory (e.g., a fragment of arithmetic). In collaboration with Paula Chocron (U. Buenos Aires, former intern in Cassis) and Pascal Fontaine (project-team Veridis), we have investigated both cases.

The notion of gentle theory has been introduced in the last few years as one solution to go beyond the restriction of stable infiniteness, but in the case of disjoint theories. In [36], [59], we adapt the notion of gentle theory to the non-disjoint combination of theories sharing only unary predicates (plus constants and the equality). Like in the disjoint case, combining two theories, one of them being gentle, requires some minor assumptions on the other one. We show that major classes of theories, i.e., Loewenheim and Bernays-Schoenfinkel-Ramsey, satisfy the appropriate notion of gentleness introduced for this particular non-disjoint combination framework.

We have also considered particular non-disjoint unions of theories connected via bridging functions [37]. We present a combination procedure which is proved correct for the theory of absolutely free data structures. We consider the problem of adapting the combination procedure to get a satisfiability procedure for the standard interpretations of the data structure.

6.2.2. Unification Modulo Equational Theories of Cryptographic Primitives

Participants: Christophe Ringeissen, Michaël Rusinowitch.

Asymmetric unification is a new paradigm for unification modulo theories that introduces irreducibility constraints on one side of a unification problem. It has important applications in symbolic cryptographic protocol analysis, for which it is often necessary to put irreducibility constraints on portions of a state. However many facets of asymmetric unification that are of particular interest, including its behavior under combinations of disjoint theories, remain poorly understood. In [42], [63] we give a new formulation of the method for unification in the combination of disjoint equational theories developed by Baader and Schulz that both gives additional insights into the disjoint combination problem in general, and furthermore allows us to extend the method to asymmetric unification, giving the first unification method for asymmetric unification in the combination of disjoint theories.

Some attacks exploit in a clever way the interaction between protocol rules and algebraic properties of cryptographic operators. In [79], we provide a list of such properties and attacks as well as existing formal approaches for analyzing cryptographic protocols under algebraic properties.

We have further investigated unification problems related to the Cipher Block Chaining (CBC) mode of encryption. We first model chaining in terms of a simple, convergent, rewrite system over a signature with two disjoint sorts: list and element. The 2-sorted convergent rewrite system is then extended into one that captures a block chaining encryption-decryption mode at an abstract level, (using no AC-symbols); unification modulo this extended system is shown to be decidable [13].

6.2.3. Enumeration of Planar Proof Terms

Participant: Alain Giorgetti.

By the Curry-Howard isomorphism, simply typed lambda terms correspond to natural deduction proofs in minimal logic. Abramsky has introduced a notion of planarity for proof terms, from a graphical representation of proofs. Noam Zeilberger and Alain Giorgetti have initiated an enumerative study of normal planar lambda terms. At the MAP 2014 workshop in Paris, Noam Zeilberger conjectured that the sequence counting the number of closed normal planar lambda terms by increasing size may coincide with the one counting the number of rooted planar maps by number of edges. Zeilberger and Giorgetti started discussing this curious coincidence at the workshop and found a proof that both families are in size-preserving bijection [70]. Although the formal aspect is not emphasized in the paper, the use of formal representations of both normal planar lambda terms and rooted planar maps, of logic programming and a proof assistant software helped much in more quickly finding the bijection. Moreover the result puts a new light on the structure of proofs in minimal logic.

6.3. Security Protocol Verification

The design of cryptographic protocols is error-prone. Without a careful analysis, subtle flaws may be discovered several years after the publication of a protocol, yielding potential harmful attacks. In this context, formal methods have proved their interest for obtaining good security guarantees. Many analysis techniques have been proposed in the literature [76]. We have edited a book [71] where each chapter presents an important and now standard analysis technique. This year, we have written a tutorial that may serve when teaching formal analysis of security protocols [26]. We develop new techniques for richer primitives, wider classes of protocols and higher security guarantees. In Section 6.5.3 we consider derived testing techniques for verifying protocol implementations.

6.3.1. Voting Protocols

Participants: Véronique Cortier, David Galindo-Chacon, Stéphane Glondu, Steve Kremer.

Voting is a cornerstone of democracy and many voting systems have been proposed so far, from old paper ballot systems to purely electronic voting schemes. Although many works have been dedicated to standard protocols, very few address the challenging class of voting protocols.

One famous e-voting protocol is Helios, an open-source web-based end-to-end verifiable electronic voting system, used e.g., by UCL and the IACR association in real elections. One main advantage of Helios is its verifiability, up-to the ballot box (a dishonest ballot box may add ballots). We have defined a variant of Helios, named Belenios, that prevents from ballot stuffing, even against a dishonest ballot box. Our approach consists in introducing an additional authority that provides credentials that the ballot box can verify but not forge. Ballot privacy of Belenios then follows from ballot privacy of Helios. For full verifiability, we had first to adapt existing definitions of verifiability in the case of a corrupted ballot box and then prove verifiability of Helios [40], [61].

This new version has been implemented by Stéphane Glondu and has been tested in an election that involved the members of the Inria Nancy-Grand Est center and the LORIA lab (about 500 people that had to chose the best LORIA pictures).

Even a basic property like ballot secrecy is difficult to define formally and several definitions co-exist. We studied all game-based privacy definitions of the literature and discovered that none of them was satisfactory: they were either limited (not fully modeling e-voting protocols), or too strong (incompatible with verifiability), or even flawed for a few of them. Based on our findings, we have proposed a new game-based privacy definition BPRIV, proved that it implies simulation-based privacy and showed that it is realized by the Helios protocol.

Existing automated analysis techniques are inadequate to deal with commonly used cryptographic primitives, such as homomorphic encryption and mix-nets, as well as some fundamental security properties, such as verifiability. In collaboration with Matteo Maffei and Fabienne Eigner (Saarland University) we propose a novel approach based on refinement type systems for the automated analysis of two fundamental properties of e-voting protocols, namely, vote privacy and verifiability. We demonstrate the effectiveness of our approach by developing the first automated analysis of Helios using an off-the-shelf type-checker.

We have presented some of our results on e-voting as plenary speaker of FLOC 2014 [31].

6.3.2. Other Families of Protocols

Participants: Véronique Cortier, Steve Kremer, Cyrille Wiedling.

Securing routing Protocols. The goal of routing protocols is to construct valid routes between distant nodes in the network. If no security is used, it is possible for an attacker to disorganize the network by maliciously interacting with the routing protocols, yielding invalid routes to be built. We have proposed a new model and an associated decision procedure to check whether a routing protocol can ensure that honest nodes only accept valid routes, even if one of the nodes of the network is compromised. This result has been obtained for a bounded number of sessions, adapting constraint solving techniques to node topologies as well as some families of recursive tests, used in routing protocols [15].

Security APIs. In some systems, it is not possible to trust the host machine on which sensitive codes are executed. In that case, security-critical fragments of a program should be executed on some tamper resistant device (TRD), such as a smartcard, USB security token or hardware security module (HSM). The exchanges between the trusted and the untrusted infrastructures are ensured by special kind of API (Application Programming Interface), that are called *security APIs*. We have designed a generic API for key-management based on key hierarchy [20], that can self-recover from corruption of arbitrary keys, provided the number of corrupted, active keys is smaller than some threshold.

Security APIs, key servers and protocols that need to keep the status of transactions, require to maintain a global, non-monotonic state, e.g., in the form of a database or register. However, most existing automated verification tools do not support the analysis of such stateful security protocols - sometimes because of fundamental reasons, such as the encoding of the protocol as Horn clauses, which are inherently monotonic. A notable exception is the recent tamarin prover which allows specifying protocols as multiset rewrite (MSR) rules, a formalism expressive enough to encode states. As multiset rewriting is a “low-level” specification language with no direct support for concurrent message passing, encoding protocols correctly is a difficult and error-prone process. In [45] we propose a process calculus with constructs for manipulation of a global state by processes running in parallel. We show that this language can be translated to MSR rules whilst preserving all security properties expressible in a dedicated first-order logic for security properties. The translation has been implemented in a prototype tool which uses the tamarin prover as a backend. We apply the tool to several case studies among which a simplified fragment of PKCS#11, the Yubikey security token, and an optimistic contract signing protocol.

6.3.3. Automated Verification of Indistinguishability Properties

Participants: Vincent Cheval, Rémy Chrétien, Véronique Cortier, Steve Kremer.

New emerging classes of protocols such as voting protocols often require to model less classical security properties, such as anonymity properties, strong versions of confidentiality and resistance to offline guessing attacks. Many of these properties can be modelled using the notion of indistinguishability by an adversary, which can be conveniently modeled using process equivalences.

Active case, unbounded number of sessions. We have studied how to reduce the search space for attacks on equivalence-based properties, for an unbounded number of sessions. Specifically, we have shown [38], [60] that if there is an attack then there is one that is well-typed. Our result holds for a large class of typing systems and a large class of *determinate* security protocols. Assuming finitely many nonces and keys, we can derive from this result that trace equivalence is decidable for an unbounded number of sessions for a class of tagged protocols, yielding one of the first decidability results for the unbounded case. As an intermediate result, we also provide a novel decision procedure in the case of a bounded number of sessions.

Active case, bounded number of sessions. We previously proposed a procedure for approximating trace equivalence in the case of a bounded number of sessions, i.e., for a replication free fragment of a cryptographic process calculus. The procedure is implemented in the *Akiss* tool. While we proved soundness and correctness for any convergent rewrite system that has the finite variant property, termination of the procedure was still an open question. We have recently shown that the procedure indeed terminates for the class of subterm convergent rewrite systems. The submission of this result is in preparation.

6.3.4. Securely Composing Protocols

Participants: Véronique Cortier, Steve Kremer, Éric Le Morvan.

Protocols may interact with an arbitrary attacker which yields a verification problem that has several sources of unboundedness (size of messages, number of sessions, etc.). In [14], we characterise a class of protocols for which deciding security for an unbounded number of sessions is decidable, by the means of a composition result. More precisely, we present a simple transformation which maps a protocol that is secure for a bounded number of protocol sessions (a decidable problem) to a protocol that is secure for an unbounded number of sessions. The precise number of sessions that need to be considered is a function of the security property and we show that for several classical security properties a single session is sufficient. Therefore, in many cases our result yields a design strategy for security protocols: (i) design a protocol intended to be secure for a single session; and (ii) apply our transformation to obtain a protocol which is secure for an unbounded number of sessions.

Protocols are often built in a modular way. For example, authentication protocols may assume pre-distributed keys or may assume secure channels. However, when an authentication protocol has been proved secure assuming pre-distributed keys, there is absolutely no guarantee that it remains secure when executing a real protocol for distributing the keys. How the security of these protocols can be combined is an important issue that is studied in the PhD thesis started by Éric Le Morvan.

6.3.5. Soundness of the Dolev-Yao Model

Participants: Véronique Cortier, Guillaume Scerri.

All the previous results rely on symbolic models of protocol executions in which cryptographic primitives are abstracted by symbolic expressions. This approach enables significantly simple and often automated proofs. However, the guarantees that it offers have been quite unclear compared to cryptographic models that consider issues of complexity and probability. A somewhat recent line of research consists in identifying cases where it is possible to obtain the best of both cryptographic and formal worlds: fully automated proofs and strong, clear security guarantees.

Gergei Bana and Hubert Comon have proposed a new framework [73] where the symbolic model now specifies what an attacker *cannot* do instead of specifying what it can do. Checking protocols security can then be reduced to checking inconsistency of some set of first order formula. During his PhD, Guillaume Scerri studies how to develop a (polynomial) decision procedure for deciding consistency of sets of formulas, for some class of formulas corresponding to security protocols. This procedure has been extended and implemented, yielding the tool SCARY that can successfully analyse several protocols of the literature [52].

6.3.6. Advanced Cryptographic Models

Participant: David Galindo-Chacon.

A classical approach in cryptographic research consists in weakening the assumptions cryptographic primitives are built upon. The following works belong to this research line.

We generalize the decisional problem that was used to prove the security of a well-known hierarchical identity-based encryption scheme by Boneh, Boyen and Goh. We argue that our new problem is strictly harder than the original problem, and thus the security of the aforementioned cryptographic primitive is laid on even stronger foundations [24].

It is known how to transform certain canonical three-pass identification schemes into signature schemes via the Fiat-Shamir transform. Pointcheval and Stern showed that those schemes are existentially unforgeable in the random-oracle model leveraging the, at that time, novel forking lemma. Recently, a number of 5-pass identification protocols have been proposed. Extending the above technique to capture 5-pass identification schemes would allow to obtain novel unforgeable signature schemes. In this paper, we provide an extension of the forking lemma (and the Fiat-Shamir transform) in order to assess the security of what we call n-generic signature schemes. These include signature schemes that are derived from certain $(2n+1)$-pass identification schemes. In doing so, we put forward a generic methodology for proving the security of a number of signature schemes derived from $(2n+1)$-pass identification schemes for $n \geq 2$. As an application of this methodology, we obtain two new code-based existentially-unforgeable signature schemes, along with a security reduction. In particular, we solve an open problem in multivariate cryptography posed by Sakumoto, Shirai and Hiwatari at CRYPTO 2011 [22].

Traditionally, symbolic and computational models for cryptographic protocols do not take into account the data leaked due to the physical nature of the cryptographic computations. Recently, the research area of leakage-resilient cryptography has emerged in order to cope with this source of attacks in the computational model. We have studied a conjecture that states that an ElGamal-based public-key encryption scheme with stateful decryption resists lunch-time chosen ciphertext and leakage attacks in the only computation leaks information model. We have given a non-trivial upper bound on the amount of leakage tolerated by this conjecture. More precisely, we prove that the conjecture does not hold if more than a $(\frac{3}{8}+o(1))$ fraction of the bits are leaked at every decryption step, by showing a lunch-time attack that recovers the full secret key. The attack uses a new variant of the Hidden Number Problem, that we call Hidden Shares - Hidden Number Problem, which is of independent interest [25].

6.4. Model-based Verification

We have investigated extensions of regular model-checking to new classes of rewrite relations on terms. We have studied specification and proof of modular imperative programs, as well as of modal workflows.

6.4.1. Tree Automata with Constraints

Participants: Pierre-Cyrille Héam, Olga Kouchnarenko.

Tree automata with constraints are widely used to tackle data base algorithmic problems, particularly to analyse queries over XML documents. The model of Tree Automata with Global Constraints (TAGED) is a model introduced in 2009 for these purposes. The membership problem for TAGED is known to be NP-complete. The emptiness problem for TAGED is known to be decidable and the best known algorithm in the general case is non elementary. In collaboration with Vincent Hugot, we show that if there is at least one negative constraint, the problem is already NP-hard [64]. In the future, we plan to investigate upper bounds for the emptiness problem with a unique negative constraint. We also plan to study the complexity of the universality problem with a single constraint.

6.4.2. Random Generation of Finite Automata

Participant: Pierre-Cyrille Héam.

Developing new algorithms and heuristics raises crucial evaluation issues, as improved worst-case complexity upper-bounds do not always transcribe into clear practical gains. A suite for software performance evaluation can usually gather three types of entries: benchmarks, hard instance and random inputs, that deliver average complexity estimations, for which the catch resides in obtaining a meaningful random distribution (for instance a uniform random distribution).

In collaboration with Jean-Luc Joly, we investigate the problem of randomly and uniformly generating deterministic pushdown automata [65]. Based on a recursive counting approach, we propose a polynomial time algorithm for this purpose. The influence of the accepting condition on the generated automata is also experimentally studied.

Partially ordered automata are finite automata where simple loops have length one. They appear in several verification techniques, such as computing closures under semi-commutation relations or studying FIFO systems. In [68], we use a Markov chain based approach to randomly - and uniformly - generate deterministic partially ordered automata. The advantage of such a technique is its flexibility, allowing for instance to easily bound the number of loops. Experiments show that the mixing time seems to be polynomial, providing a tractable approach.

6.4.3. Verification of Linear Temporal Patterns over Finite and Infinite Traces

Participants: Pierre-Cyrille Héam, Olga Kouchnarenko.

In the regular model-checking framework, reachability analysis can be guided by temporal logic properties, for instance to achieve the counter example guided abstraction refinement (CEGAR) objectives. A way to perform this analysis is to translate a temporal logic formula expressed on maximal rewriting words into a “rewrite proposition” – a propositional formula whose atoms are language comparisons, and then to generate semi-decision procedures based on (approximations of) the rewrite proposition. In collaboration with Vincent Hugot, we have investigated suitable semantics for LTL on maximal rewriting words and their influence on the feasibility of a translation, and we have proposed a general scheme providing exact results for a fragment of LTL corresponding mainly to safety formulæ, and approximations for a larger fragment.

6.4.4. Machine-Learning Techniques for Regular Model-Checking

Participants: Maxime Bride, Pierre-Cyrille Héam.

Using a machine-learning approach, we address the general problem of regular model-checking of computing $R^*(L)$, when L is a regular language and R a relation. Rather than developing specific algorithms to compute $R^*(L)$, it consists in using Angluin style’s algorithms. In [58], we focus on the generation of examples, counter-examples and on the design of an oracle for the specific case of semi-commutation relations. Experiments are promising, particularly for the sizes of the obtained automata, which are quite smaller than with dedicated algorithms.

6.4.5. Constraint Solving for Verifying Modal Workflow Specifications

Participants: Hadrien Bride, Olga Kouchnarenko.

Workflow Petri nets are well suited for modelling and analysing discrete event systems exhibiting behaviours such as concurrency, conflict, and causal dependency between events. They represent finite or infinite-state processes, and several important verification problems, like reachability or soundness, are known to be decidable. Modal specifications introduced in [84] allow loose or partial specifications in a framework based on process algebras.

Our work in [34] focuses on the verification of modal workflow specifications using constraint solving as a computational tool. Its main contribution consists of a formal framework based on constraint systems to model executions of workflow Petri nets and their structural properties, as well as to verify their modal specifications. An implementation and promising experimental results obtained within the proposed approach constitute a practical contribution. In particular, a business process example from the IT domain enables to successfully assess the reliability of our contributions.

6.4.6. Rewriting-based Mathematical Model Transformations

Participants: Walid Belkhir, Alain Giorgetti.

Since 2011 we collaborate with the Department "Temps-Fréquence" of the FEMTO-ST institute (Franche-Comté Electronique Mécanique Thermique et Optique - Sciences et Technologies, CNRS UMR 6174) on the formalization of asymptotic methods (based on two-scale convergence). The goal is to design a software, called *MEMSALab*, for the automatic derivation of multiscale models of arrays of micro- and nanosystems. In this domain a model is a partial differential equation. Multiscale methods approximate it by another partial differential equation which can be numerically simulated in a reasonable time. The challenge consists in taking into account a wide range of geometries combining thin and periodic structures with the possibility of multiple nested scales. We have designed a transformation language facilitating the design of MEMSALab [17]. It is proposed as a Maple^{TM} package for rule-based programming, rewriting strategies and their combination with standard Maple^{TM} code. We illustrate the practical interest of this language by using it to encode two examples of multiscale derivations, namely the two-scale limit of the derivative operator and the two-scale model of the stationary heat equation. A more general framework for the derivation of the multi-scale models was established in [29].

6.5. Model-based Testing

Our research in Model-Based Testing (MBT) aims to extend the coverage of tests. The coverage refers to several artefacts: model, test scenario/property, and code of the program under test [55]. The test generation uses various underlying techniques such as symbolic animation of models [80], or symbolic execution of programs by means of dedicated constraints, SMT solvers, or model-checkers.

6.5.1. Automated Test Generation from Behavioral Models

Participants: Fabrice Bouquet, Kalou Cabrera, Jérome Cantenot, Frédéric Dadeau, Jean-Marie Gauthier, Julien Lorrain, Alexandre Vernotte.

We have developed an original model-based testing approach that takes a behavioral view (modelled in UML) of the system under test and automatically generates test cases and executable test scripts according to model coverage criteria [18]. We continue to extend this result to SysML specifications for validating embedded systems. We apply this method on smartSurface [44].

In the context of the FSN DAST project on Dynamic Application Security Testing, we investigated the use of a model-based testing approach for vulnerability testing in web applications. We designed a process based on two artefacts. First, a generic UML model, that is used to represent the web application entities (pages, forms, etc.), coupled with OCL constraints that describe the business logics of the application. Second, a set of test purposes, that will look for specific vulnerabilities (cross-site scripting, SQL injections, etc.). We have implemented a research prototype and applied it on several case studies. It has shown its effectiveness to detect vulnerabilities on already deployed web applications [50].

6.5.2. Scenario-Based Verification and Validation

Participants: Fabrice Bouquet, Kalou Cabrera, Frédéric Dadeau.

Test scenarios represent an abstract test case specification that aims at guiding the model animation in order to produce relevant test cases. Contrary to the previous section, this technique is not fully automated since it requires the user to design the scenario, in addition to the model.

We have proposed a dedicated formalism to express test properties. A test property is first translated into a finite state automaton which describes a monitor of its behaviors. We have also proposed dedicated property coverage criteria that can be used either to measure the property coverage of a given test suite, or to generate test cases, exercising nominal or robustness aspects of the property [41]. This process has been fully tool-supported into an integrated software prototype[0]. This process has been designed during the ANR TASCCC project (2009-2012) and was continued during the ANR ASTRID OSEP project (2012-2013). The industrialization of this approach, and its integration within commercial test generation tools has started with the ANR ASTRID Maturation MBT_Sec project (2014-2015).

In the context of the SecureChange project, we have also investigated the evolution of test scenarios. As the system evolves, the model evolves, and the associated test scenarios may also evolve. We are currently extending the test generation and management of system evolutions to ensure the preservation of the security [43].

6.5.3. Mutation-based Testing of Security Protocols

Participants: Frédéric Dadeau, Pierre-Cyrille Héam, Ghazi Maatoug, Michaël Rusinowitch.

We have proposed a model-based penetration testing approach for security protocols [41]. This technique relies on the use of mutations of an original protocol, proved to be correct, for injecting realistic errors that may occur during the protocol implementation (e.g., re-use of existing keys, partial checking of received messages, incorrect formatting of sent messages, use of exponential/xor encryption, etc.). Mutations that lead to security flaws are used to build test cases, which are defined as a sequence of messages representing the behavior of the intruder. We have applied our technique on protocols designed in HLPSL, and implemented the protocol mutation tool jMuHLPSL that performs the mutations. The mutants are then analyzed by *CL-AtSe*. We have experimented our approach on a set of protocols, and we have shown the relevance of the proposed mutation operators and the efficiency of the *CL-AtSe* to conclude on the vulnerability of a protocol and produce an attack trace that can be used as a test case for implementations. We applied our approach on the Paypal Express protocol, and we were able to retrieve an existing attack trace on this protocol[0]. We also investigated the transformation of an attack trace into executable tests scripts. To achieve that, we have proposed to automatically generate skeletons of Java test programs that the validation engineer only has to fill in order to concretize the steps of the test. Experimentations on these principles have been described in [53].

6.5.4. Code and Contract-based Test Generation and Static Analysis

Participants: Fabrice Bouquet, Frédéric Dadeau, Ivan Enderlin, Alain Giorgetti.

With the CEA we have developed a test generation technique based on C code and formal specifications, to facilitate deductive verification, in a new tool named StaDy [67], [49], [51]. The tool integrates the concolic test generator PathCrawler within the static analysis platform Frama-C. StaDy is able to handle the ANSI C Specification Language (ACSL) of the framework and other Frama-C plug-ins are able to reuse results from the test generator. This tool is designed to be the foundation stone of modular static and dynamic analysis combinations in the Frama-C platform.

We have designed a new annotation language for PHP, named PRASPEL (for *PHP Realistic Annotation SPEcification Language*). This language relies on *realistic domains* which serve two purposes. First, they assign to a data a domain that is supposed to be specific w.r.t. a context in which it is employed. Second, they provide two features that are used for test generation: (i) *samplability* makes it possible to automatically generate a value that belongs to the realistic domain so as to generate test data, (ii) *predicability* makes it possible to check if the value belongs to a realistic domain. This approach is tool-supported in a dedicated framework for PHP which makes it possible to produce unit test cases using random data generators, execute the test cases on an instrumented implementation, and decide the conformance of the code w.r.t. the annotations by runtime assertion checking. This principle has been extended to generate grammar-based textual data based

[0]A video of the prototype is available at: http://vimeo.com/53210102
[0]http://www.nbs-system.com/blog/faille-securite-magento-paypal.html

on various strategies, namely uniform random generation, bounded exhaustive generation and rule-coverage-based test generation. In a recent work, we have proposed a dedicated constraint solver for PHP arrays aiming to avoid rejection during the generation of array structures. Finally, we have proposed dedicated specification coverage criteria to drive the test generation process. These coverage criteria focus on the selection of a subset of a method's contract, or the selection of specific predicates or realistic domains inside the contract. The whole approach has been implemented into a dedicated framework [62] integrated with state-of-the-practice test execution environments, such as atoum.

6.5.5. Random Testing

Participants: Aloïs Dreyfus, Pierre-Cyrille Héam, Olga Kouchnarenko.

The random testing paradigm represents a quite simple and tractable software assessment method for various testing approaches. When performing random testing, the random sampler is supposed to be independent of tester choices or convictions: a solution is to exploit uniform random generators.

In [82] a method is proposed for drawing paths in finite graphs uniformly, and it is explained how to use these techniques for testing C programs within a control flow graph based approach. Nevertheless, as finite graphs often provide strong abstractions of the systems under test, many abstract tests generated by the approach cannot be played on the implementation. In [83], we have proposed a new approach, extending [82], to manage stack-call during the random test generation while preserving uniformity. In [23], we go further by investigating a way to biase the random testing, in order to optimize the probability to fulfil a coverage criterion. The new approaches have been implemented in a prototype and experimented on several examples.

6.6. Verification of Collaborative Systems

We investigate security problems occurring in decentralized systems. We develop general techniques to enforce read and update policies for controlling access to XML documents based on recursive DTDs (Document Type Definition). Moreover, we provide a necessary and sufficient condition for undoing safely replicated objects in order to enforce access control policies in an optimistic way.

6.6.1. Automatic Analysis of Web Services Security

Participants: Walid Belkhir, Michaël Rusinowitch, Mathieu Turuani, Laurent Vigneron.

Automatic composition of web services is a challenging task. Many works have considered simplified automata models that abstract away from the structure of messages exchanged by the services. For the domain of secured services (using e.g., digital signing or timestamping) we propose a novel approach to automated orchestration of services under security constraints. Given a community of services and a goal service, we reduce the problem of generating a mediator between a client and a service community to a security problem where an intruder should intercept and redirect messages from the service community and a client service till reaching a satisfying state. This orchestration specification is expressed in ASLan language, a formal language designed for modeling Web Services tied with security policies that was developed in AVANTSSAR project. The AVANTSSAR Orchestrator (presented in [56]) generates an attack trace describing the execution of the mediator and translates it into ASLan. Then we can check with automatic tools that this ASLan specification verifies required security properties such as secrecy and authentication. If no flaw is found, we can compile the ASLan specification into a Java servlet that can be used to execute the orchestration.

In [16] we develop our alternative approach based on *parametrized automata*, a natural extension of finite-state automata over infinite alphabet. In this model the transitions are labeled with constants or variables that can be refreshed in some specified states. We prove several closure properties for this class of automata and study their decision problems. We show the applicability of our model to Web services handling data from an infinite domain. We introduce a notion of simulation that enables us to reduce the Web service composition problem to the construction of a simulation of a target service by the asynchronous product of existing services, and prove that this construction is computable. The existence of a service orchestrator solving a service composition problem can alternatively be reduced to the satisfiability of formula in parametrized propositional dynamic logic, and the latter was shown decidable in [33].

We now work on synthesizing composed services that satisfy required security policies.

6.6.2. Secure Querying and Updating of XML Data

Participants: Abdessamad Imine, Houari Mahfoud, Michaël Rusinowitch.

It is increasingly common to find XML views used to enforce access control as found in many applications and commercial database systems. To overcome the overhead of view materialization and maintenance, XML views are necessarily virtual. With this comes the need for answering XML queries posed over virtual views, by rewriting them into equivalent queries on the underlying documents. A major concern here is that query rewriting for recursive XML views is still an open problem, and proposed approaches deal only with non-recursive XML views. Moreover, a small number of works have studied the access rights for updates. In [11], we present SVMAX (Secure and Valid MAnipulation of XML), the first system that supports specification and enforcement of both read and update access policies over arbitrary XML views (recursive or non). SVMAX defines general and expressive models for controlling access to XML data using significant class of XPath queries and in the presence of the update primitives of W3C XQuery Update Facility. Furthermore, SVMAX features an additional module enabling efficient validation of XML documents after primitive updates of XQuery. The wide use of W3C standards makes of SVMAX a useful system that can be easily integrated within commercial database systems. We give extensive experimental results, based on real-life DTDs, that show the efficiency and scalability of our system.

6.6.3. Secure Computation in Social Networks

Participants: Bao Thien Hoang, Abdessamad Imine, Huu Hiep Nguyen, Michaël Rusinowitch.

Online social networks are currently experiencing a peak and they resemble real platforms of social conversion and content delivery. Indeed, they are exploited in many ways: from conducting public opinion polls about any political issue to publish social graph data for achieving in-depth studies. To securely perform these large-scale computations, we need the design of reliable protocols to ensure the data privacy. To address the polling problem in social networks (where the privacy of exchanged information and user reputation are very critical), we provide a simple decentralized polling protocol that relies on the current state of social graphs. More explicitly, we define one family of social graphs that satisfy what we call the m-broadcasting property (where m is less than or equal to a minimum node degree). We show their structures enable low communication cost and constitute necessary and sufficient condition to ensure vote privacy and limit the impact of dishonest users on the accuracy of the polling output. To securely publish social graph data, we focus on the problem of anonymizing a deterministic graph by converting it into an uncertain form [48], [47]. We first analyze drawbacks in a recent uncertainty-based anonymization scheme and then propose Maximum Variance, a novel approach that gains better tradeoff between privacy and utility. Towards a fair comparison between the anonymization schemes on graphs, the second contribution of our work is to describe a quantifying framework for graph anonymization by assessing privacy and utility scores of typical schemes in a unified space.

6.6.4. Safe and Secure Protocols for Collaborative Applications

Participants: Abdessamad Imine, Michaël Rusinowitch.

The Operational Transformation (OT) approach, used in many collaborative editors, allows a group of users to concurrently update replicas of a shared object and exchange their updates in any order. The basic idea is to transform any received update operation before its execution on a replica of the object. Designing transformation functions for achieving convergence of object replicas is a critical and challenging issue. In this work, we investigate the existence of transformation functions [27]. From the theoretical point of view, two properties, named TP1 and TP2, are necessary and sufficient to ensure convergence. Using controller synthesis technique, we show that there are some transformation functions, which satisfy TP1 for the basic signatures of insert and delete operations. But, there is no transformation function, which satisfies both TP1 and TP2. Consequently, a transformation function which satisfies both TP1 and TP2 must necessarily have additional parameters in the signatures of some update operations. Accordingly, we provide a new transformation function and show formally that it ensures convergence.

In [19], we propose a generic access control model based on replicating the shared document and its authorization policy at the local memory of each user. We consider the propagation of authorizations and their interactions. We use an optimistic approach to enforce access control in existing collaborative editing solutions in the sense that the access control policy can be temporarily violated. To enforce the policy, we resort to the selective undo approach in order to eliminate the effect of illegal document updates. To validate our approach, we implement an optimistic access control on the top of a collaboration prototype and measure its performance in the distributed grid GRID'5000 to highlight the scalability of our solution.

However, verifying whether the combination of access control and coordination protocols preserves the data consistency is a hard task since it requires examining a large number of situations. In [30], we specify this access control protocol in the first-order relational logic with Alloy, and we verify that it preserves the correctness of the system on which it is deployed, namely that the access control policy is enforced identically at all participating user sites and, accordingly, the data consistency remains still maintained.

CASTOR Project-Team

6. New Results

6.1. Physics

6.1.1. Physical studies

6.1.1.1. Parallel Kelvin-Helmholtz instability

Participants: Hervé Guillard, Marco Bilanceri, Céline Colin [CEA], Philippe Ghendrih [CEA], Giorgio Giorgiani, Boniface Nkonga, Frédéric Schwander [M2P2, AMU], Eric Serre [M2P2, AMU], Patrick Tamain [CEA].

In the scrape-off layer (SOL) of tokamaks, the flow acceleration due to the presence of limiter or divertor plates rises the plasma velocity in a sonic regime. These high velocities imply the presence of a strong shear between the SOL and the core of the plasma that can possibly trigger some parallel shear flow instability. The existence of these instabilities, denoted as parallel Kelvin-Helmholtz instability in some works have been investigated theoretically in [51] using a minimal model of electrostatic turbulence composed of a mass density and parallel velocity equations. This work showed that the edge plasma around limiters might indeed be unstable to this type of parallel shear flow instabilities. In this work, begun in 2013, we have performed large scale 3D simulations using the PlaTo platform of the same simple mathematical model to investigate this question. The numerical results confirm that in agreement with the theoretical expectations as well as with other numerical methods, the sheared flows in the SOL are subject to parallel Kelvin-Helmholtz instabilities. However, the growth rate of these instabilities is low and these computations require both a sufficient spatial resolution and a long simulation time. This makes the simulation of parallel Kelvin-Helmholtz instabilities a demanding benchmark but it also allows us to validate the parallel implementation of the PlaTo platform up to up to $\mathcal{O}(1000)$ CPU [14].

6.2. Numerical developments

6.2.1. Numerical developments

6.2.1.1. Conformal hexahedral mesh coarsening by agglomeration

Participants: Hervé Guillard, Pierre Cargemel, Youssef Mesri [IFPEN].

This work has been realized in the framework of a PhD contract with IFPEN that aims to produce a coarsening software for hex-dominant meshes. Reservoir simulation involves to compute dynamic flow of different phases in a porous medium. The initial state of the reservoir is usually pre-computed via geo-statistics methods extrapolating measures of the terrain. Therefore, the input of reservoir simulation is given as a very fine mesh containing heterogeneous data and numerical simulation on this fine mesh is usually non-practical. This work is therefore devoted to the study of an agglomeration strategy, to dynamically coarsen this fine hex-dominant mesh. The adaptivity may be driven by physics and/or geometric estimators. Ideally, the coarsening should be applied locally in low gradient regions, whereas high gradient regions keep the fine mesh. This work has been presented in the 23rd International meshing Roundtable [26]. The planned sequel of this work consists to use the notion of Central Voronoi Tesselation (CVT) to treat the regions where the mesh is not structured and to apply this strategy in different physical contexts from plasma physics to petroleum engineering.

6.2.1.2. Mapped Fourier Methods for stiff problems in toroidal geometry

Participant: Hervé Guillard.

Due to the particular geometry of tokamaks, a lot of numerical codes developed for their numerical simulations use Fourier methods. Fourier spectral or pseudo-spectral methods are extremely efficient for periodic problems. However this efficiency is lost if the solutions have zones of rapid variations or internal layers. For these cases, a large number of Fourier modes are required and this makes the Fourier method unpractical in many cases. This work investigates the use of mapped Fourier method as a way to circumvent this problem. Mapped Fourier method uses instead of the usual Fourier interpolant the composition of the Fourier interpolant with a mapping in such a way that in the computational space, the functions to represent are not stiff. This work gives some examples of the usefulness of this method and apply it to a simple model of pellet injection in tokamaks as an example of its potential interest for complex multi dimensional problem [34].

6.2.1.3. Multislope MUSCL method for general unstructured meshes

Participants: Hervé Guillard, Clément Le touze [ONERA], Angelo Murrone [ONERA].

To increase the accuracy in finite volume method, the concept of MUSCL reconstruction has been introduced in the pioneering work of van Leer in the 70'. This technique is still one of the most efficient to deal with the existence of discontinuous solutions in numerical simulations. In the MUSCL technique, a discontinuous linear approximation of the solution is reconstructed on each control volume. The main approximation problem of this method is therefore to reconstruct the slope of the solution.

The multislope concept has been recently introduced in the literature to deal with MUSCL reconstructions on triangular and tetrahedral unstructured meshes in the finite volume cell-centered context. Dedicated scalar slopes are used to compute the interpolations on each face of a given element, in opposition to the monoslope methods in which a unique limited gradient is used. The multislope approach reveals less expensive and potentially more accurate than the classical gradient techniques. Besides, it may also help the robustness when dealing with hyperbolic systems involving complex solutions, with large discontinuities and high density ratios. In this work, we have designed a generalized multislope MUSCL method for cell-centered finite volume discretizations. The method is freed from constraints on the mesh topology, thereby operating on completely general unstructured meshes. Moreover optimal second-order accuracy is reached at the faces centroids and the scheme is L^∞ stable. Special attention has also been paid to equip the reconstruction procedure with well-adapted dedicated limiters, potentially CFL-dependent. We have shown in [18] the ability of the method to deal with completely general meshes, while exhibiting second-order accuracy.

6.2.1.4. Development of a two temperature model

Participants: Hervé Guillard, Afeintou Sangam, Elise Estibals.

A two temperature (ions - electrons) model for non-magnetized plasma has been designed. The numerical scheme is a finite volume method with an approximate Riemann solver using the total energy equation and the electron entropy as main variables. This Riemann solver has been validated against standard shock tube problems and incorporated in the PlaTo platform. The solver has been implementated in toroidal geometry and tested successfully on realistic particular flows encountered in this context. The development of a reduced MHD model based on this two temperature scheme is currently studied.

6.2.1.5. Entropy Preserving Schemes for Conservation Laws

Participants: Christophe Berthon [University of Nantes], Bruno Dubroca [CEA/DAM/CESTA and University of Bordeaux 1], Afeintou Sangam.

A relaxation-type scheme has been proposed to approximate weak solutions of Ten-Moments equations with source terms [*2*]. These equations model compressible anisotropic flows. Following the technique introduced in [44], the proposed scheme is proved to be entropy preserving.

6.2.1.6. Eurofusion WPCD: Free boundary equilibrium code and control

Participants: Cédric Boulbe, Blaise Faugeras, Jean François Artaud [IRFM CEA Cadarache], Vincent Basiuk [IRFM CEA Cadarache], Emiliano Fable [Max-Planck-Institut für Plasmaphysik, Garching], Philippe Huyn [IRFM CEA Cadarache], Eric Nardon [IRFM CEA Cadarache], Jakub Urban [IPP, Academy of Sciences of the Czech Republic, Prague].

Our team is involved in the integrated modelling WPCD (Work Package Code Development) Eurofusion. This project is the continuation of the EFDA-ITM project. The goal of WPCD is to provide a european tool for tokamak simulations. Different physical codes can be coupled using Kepler environment. Machine description and physical data have been described using CPO (Consistent Physical Objet) which are used as standardized inputs and outputs for the codes.

In this project, we participate in the coupling of a free boundary equilibrium solver, the European Transport Solver (ETS) and a plasma shape and position controller. The workflow coupling TCV hybrid Simulink controller and Cedres++ using PF circuit connections has been finalized and tested on the TCV tokamak.
A new workflow coupling Cedres++ with ETS and the TCV controller has been developed and is being tested on a TCV test case. This workflow is an evolution of the coupling CEDRES++ - ETS described in [12].
A successful benchmark between the three free boundary equilibrium codes CEDRES++([15] [47]) , FREEBIE ([43]), and SPIDER ([49]) has been done on static test cases. This activity will be continued to compare the time dependent versions of the three codes.

6.2.1.7. Optimal control for scenario optimization of discharges in tokamaks

Participants: Jacques Blum, Holger Heumann.

In this project we aim for an automatic determination of optimal voltage evolutions via an optimal control formulation based on a system of partial differential equations that describes the evolution of plasma equilibrium in a tokamak. Optimal voltage evolutions are the one that ensure that the evolution of the plasma runs through predescribed, user-defined states, defined e.g. as desired evolution of shape or position. The system of partial differential equations describing the evolution of the plasma is non-linear and we use a finite element formulation together with implicit time stepping for the discretization [15]. With this approach we end up with a large but finite dimensional optimization problem with non-linear constraints. We are using SQP (sequential quadratic programming), known to be one of the fastest methods of such problems, to solve the finite-dimensional optimization problem. The performance of SQP relies on accurate derivatives of the objective function and the constraints. The derivatives related to the free boundary, derived and implements during H. Heumann's PostDoc 2011/2012 for a static optimal control problem, appeared here again and are one of the important building blocks for treating the transient case. Both in CEDRES++ and FEEQ.M we have now the capability to solve first test cases to define optimal voltage evolutions. In contrast to the static case, where the linear algebraic systems in the SQP iteration remain reasonable small, the solution of the corresponding linear system in the transient case becomes very time-consuming, which somehow limits the applicability. We are testing variants of SQP, such as BFGS-like updates for the reduced Hessian, to see whether we could speed up and improve robustness of our calculations. Fast iterative solver for large sparse linear systems is another option that we started to investigate. Fast iterative solver for linear system in transient optimal control problems governed by partial differential equations is a very active area of research and we hope to benefit from the latest developments.

6.2.1.8. Boundary reconstruction for the WEST tokamak with VacTH

Participants: Jacques Blum, Sylvain Bremond [IRFM CEA Cadarache], Cédric Boulbe, Blaise Faugeras, Holger Heumann, Philippe Moreau [IRFM CEA Cadarache], Eric Nardon [IRFM CEA Cadarache], Remy Nouailletas [IRFM CEA Cadarache], François Saint Laurent [IRFM CEA Cadarache].

This work is under progress in collaboration with the CEA. The control of the plasma in the future WEST tokamak requires the identification of its boundary in real time during a pulse. The code VacTH under development in the team enables such an identification. Several numerical developments and experiments have been conducted in order to prepare the control of the plasma in the WEST tokamak. The equilibrium code CEDRES++, also developed in the team, is used to simulate a real plasma and to generate synthetic magnetic measurements from which the plasma boundary is reconstructed using the code VacTH. A control algorithm developed by the colleagues from the CEA then uses this knowledge of the plasma shape to adapt the currents flowing in the poloidal field coils in order to achieve a desired evolution of the plasma.

6.2.1.9. Equilibrium reconstruction for ASDEX UpGrade (AUG) with Vacth-Equinox

Participants: Blaise Faugeras, Rui Cohelo [IPFN, IST, Lisbonne], Patrick Mccarthy [National University of Ireland University College Cork].

Within the framework of the WPCD EUROFUSION the code VacTH-Equinox has been adapted to enable equilibrium reconstruction for AUG. The identification of the current density pedestal required the development of a specific regularization scheme allowing weaker regularization close to the plasma boundary and stronger close to the magnetic axis.

6.2.1.10. Taylor-Galerkin stabilized Finite Element

Participants: José Costa, Boniface Nkonga.

The theoretical part of Taylor-Galerkin/Variational multi-scales (TG/VMS) strategy applied to MHD and reduced MHD modeling has been achieved last year. The final method amounts to add in the finite element formulation, a self-adjoint operator associated to the most critical hyperbolic component of the system to be solved. The design of the critical contours and the identification of associated waves to be stabilized is problem dependent and related to the Jacobian matrix. We have focused this year on the validations of this strategy and the improvement of the linearization used for stabilization. For application to plasma configurations with X-point, we need to reconsider the consistency with equilibrium and the Bohm boundary conditions on open flux walls.

6.2.1.11. Toward full MHD numerical modeling with $\mathcal{C}^1$ FE

Participants: José Costa, Giorgio Giorgiani, Hervé Guillard, Boniface Nkonga.

In this context the single fluid full MHD model is considered and the divergence free constraint on the magnetic field is achieved by introduction of a potential vector. The use of the potential vector has the additional advantage that the toroidal component is the magnetic flux of the Grad-Shafranov equilibrium. However, using the potential vector as variable introduces higher order derivatives in the system and classical C0 finite elements cannot be directly applied. This is why our finite element strategies use shape/test functions whose derivatives have global continuity in space (smooth finite elements). The global approach uses cross product shape/test functions between poloidal(2D) and toroidal(1D). In the 2D poloidal plane, discretization uses either quadrangular or triangular elements.

This year we have focused on the numerical analysis associated to the full MHD discretization in configurations with open flux surfaces. In order to derive efficient strategies for the full MHD in the potential vector formulation, the Gauge condition on the potential vector and the boundary conditions have been enforced by penalizations. For the Gauge condition it gives rise to element contributions but also boundary integrals that should be computed on curved surfaces that sometime fitted the magnetic surfaces. Equations are formulated in semi-conservative form such as to apply integration by part. Therefore, boundary conditions can be viewed as evolution of fluxes or variables. Integral formulation on the boundary is very useful for higher order finite elements and also for easier treatment of corners. Indeed in this context the boundary conditions are edges/surfaces oriented and boundary corners are driven by the neighborhood edges penalizations. This strategy is the one that will be used for future developments.

2D Quadrangular Cubic Bezier Finite Elements: This finite element is used for a while for reduced MHD models in the software Jorek. Reduced MHD is used to project the momentum equation in a space orthogonal to the equilibrium. When full MHD models are used, the momentum equation needs to be projected in the equilibrium space and this projection should be consistent with the Grag-Shafranov equilibrium that is used to compute the initial state. This has been achieved by a proper computation of the JxB contribution in the momentum equation, taking into account the poloidal variation of the toroidal component of the magnetic field. Detailed analysis has been performed. The next year will be devoted to implementations and numerical validations.

2D Triangular Powel-Sabin Finite Elements: In order to avoid some mesh singularities when using quadrangular meshes for complex geometries and flux surfaces shapes, triangular elements are a possible option. It is not so easy to derive smooth finite element on triangle with reduced number of degree of freedom. The Bell reduced-quintic finite elements we have considered in the previous years have too much unknowns (6 per vertex). Powell-Sabin splines are piecewise quadratic polynomials with a global $\mathcal{C}^1$-continuity and 3 unknowns per vertex, they have a local support, they form a convex partition of unity, they are stable, and they have a geometrically intuitive interpretation involving control triangles. Construction of the Powel-Sabin splines needs some geometrical tools that have been developed: Minimum area enclosing triangle of a set of control points (sequential and parallel). This construction is applied to each vertex of the triangular mesh and used to derive the local shape/test functions. These Powel-Sabin splines have been used successfully in the area of computer aided geometric design for the modeling and fitting of surfaces. We have used the Powell-Sabin (PS) splines for the approximation of elliptic partial differential equations (including Grad-Shafranov) in a rectangular domain. In this context have recovered the optimal rate of convergence (order 3). Preliminary result has been obtained for hyperbolic isothermal 2D Euler equations with TG/VMS stabilization. Our aim in the coming years is to apply these PS splines to full MHD in a toroidal geometry.

6.2.1.12. Genuinely multidimensional Riemann Solver

Participants: Jeaniffer Vides, Boniface Nkonga.

Multidimensional Riemann solvers were pioneered by Abgrall. Abgrall, Maire, Nkonga, Després and Loubere have extensively developed them especially as node-solvers for Lagrangian hydrodynamics. Another strain of work comes from explorations by Wendroff and Balsara who took a space-time approach. In this work, the resolved state is obtained via space-time integration over a wave model, just as was done by Wendroff and Balsara. However, an algebraic approach is used for the development of the fluxes. It is, therefore, shown that the multidimensional fluxes can be obtained by application of jump conditions at the boundaries of the wave model. The problem is of course over determined with the result that the shock jump conditions are only satisfied approximately in a least squares sense. Even so, this work gives us new perspective on multidimensional Riemann solvers. The litteral satisfaction of the shock jump conditions (up to least squares approximation) makes it easier to understand multidimensional Riemann solvers as a natural extension of the one-dimensional Riemann solvers. Contributions have also been made on the development of a minimalist wave model, which might help in reducing dissipation. Further innovations are reported on the assembling of fluxes based on the structure of the wave model, and those innovations are potentially useful. For MHD the CT approach consists of constraining the transport of magnetic field so that the divergence is always kept zero. The method relies on exploiting the dualism between the flux components and the electric field. Since the electric field is needed at the edges of the mesh, the multidimensional Riemann solver can also provide the electric field. By running an extensive set of simulations, it is shown that the multidimensional Riemann solver is robust and can be used to obtain divergence-free formulations for MHD that perform well on several stringent calculations. Future work will improve this strategy by enriching the description of the strongly interaction of waves.

6.2.1.13. Multi scales approximations of "Shallow water" flows

Participants: Jeaniffer Vides, Boniface Nkonga.

The terminology "Shallow water" is used to characterize thin flows on curved surfaces. It is customary for this type of flows to use the incompressible Navier-Stokes equations to asymptotically derive reduced models for the evolution of the depth integrated speed and the thickness of the flow. Reduced model are mainly hyperbolic and finite volume method are often used for their numerical approximation. Approximations strategies are generally structured as follow:

- Construction of a global coordinate system associated with an assumed analytical surface.
- Reduction of the model relatively to the global coordinate system
- Approximation of the surface by a finite number of elements.
- Approximation of the reduced model using the discrete surface.

In the context of real applications, it is presumptuous to expect an analytical formulation of the surface. From the data provided by observation satellites, we can usually extract a discrete description of the surfaces that drives thin flow. Therefore, it is more practical to use the discrete description as the starting point of the resolution strategy. This is the angle of approach that we have considered. We locally define two mesh scales: the element scale and the cell scale. The discrete mapping and the reduced model are defined at the element scale and the average values that evolve in time are defined at the cell scale. First applications have been successfully performed. We will now continue your investigations and include relevant physics at each scale, including sheared flows. We will also examine the use of multi dimensional Riemann solver in this context.

6.2.1.14. Computational Magnetohydrodynamics with Discrete Differential Forms

Participants: Holger Heumann, Ralf Hiptmair [SAM, ETH Zürich, Switzerland], Cecilia Pagliantini [SAM, ETH Zürich, Switzerland].

Differential forms, or equivalently exterior calculus, are a natural framework for electromagnetics; not only for a better understanding of the theoretical foundation, but also for the development of numerical methods. Keywords are the Hodge decomposition or the de Rham complex that are at the bottom of recent development of efficient multigrid methods or stable mixed finite element methods. Thinking in terms of such co-ordinate free differential forms offers considerable benefits as regards to the construction of structure preserving spatial discretizations.

In the present project, we aim at developing a new approach for the numerical treatment of resistive magnetohydrodynamics where a Galerkin discretization of the electromagnetic part based on finite element exterior calculus (FEEC) will be coupled to advanced finite volume methods for the approximation of the balance laws for the fluid.

The latest developments involved the extension and analysis of the stablized Galerkin schemes for advection of differential forms introduced in [48] to the case of time-dependent and non-regular flow fields.

6.2.1.15. Hamilton-Jacobi Formulation for Vlasov-Poisson

Participants: Holger Heumann, Eric Sonnendrücker [IPP, Max-Planck-Institute Garching, Germany], Philip J. Morrison [Institute for Fusion Studies, Austin, USA].

The phase space mapping induced by the solution of the Vlasov-Poisson problem is a symplectic mapping (or canonical transformation in physics literature) solving Hamilton's equations. In this project we are developing numerical methods that are based on this formulation. We derived and implemented new finite difference schemes for the corresponding Hamilton-Jacobi equation, that circumvent the projection of the distribution function inherent in Lagrangian methods. First numerical results for standard test problems show the ability of increased resolution of fine-scale effects.

6.2.1.16. Entropy viscosity technique

Participants: Richard Pasquetti, Jean-Luc Guermond [Texas A & M University], Boyan Popov [Texas A & M University].

The entropy viscosity technique allows to address hyperbolic equations by introducing a strongly non linear viscous term where needed, especially at shocks. The basic idea is to set up a viscosity from the residual of the entropy inequality together with a $O(h)$ upper bound proportional to the local wave speed. In view of addressing situations where vaccum may appear in the tokamak, we have considered the shallow water equations with topography and in situations where dry-wet transitions occur. Using a RK scheme in time and a spectral element method (SEM) in space, we have proposed a variant of the entropy technique, that mainly consists of using the viscosity upper bound in the dry parts, to obtain satisfactory results. This work was presented in [30], [22] and a publication is under review.

6.2.1.17. Bohm boundary conditions

Participants: Sébastian Minjeaud, Richard Pasquetti.

In the frame of the ANR project ESPOIR, our partners have proposed a penalty method to enforce the Bohm criterion (Mach number greater than one at the tokamak plates). This approach has been justified by considering a “minimal transport model” that consists of a 1-dimensional non linear hyperbolic system of two equations, that govern the evolutions of the density and velocity. The approach and further developments are described in three recent papers published in the journal of computational physics. Considering the same hyperbolic system, we have proposed a direct way to enforce the Bohm criterion in the frame of an explicit time marching. Using a SVV stabilized SEM it is then possible to resolve the same problem with spectral accuracy. This paper is now in press and will be published as a JCP note.

6.2.1.18. A numerical scheme for fluid-particules flows

Participants: Florent Berthelin, Thierry Goudon, Sebastian Minjeaud.

We propose a numerical scheme for the simulation of fluid-particles flows with two incompressible phases. The numerical strategy is based on a finite volume discretization on staggered grids, with a flavor of kinetic schemes in the definition of the numerical fluxes. We particularly pay attention to the difficulties related to the volume conservation constraint and to the presence of a close-packing term which imposes a threshold on the volume fraction of the disperse phase. We are able to identify stability conditions on the time step to preserve this threshold and the energy dissipation of the original model. The numerical scheme is validated with the simulation of sedimentation flows.

6.2.1.19. Identification and forecast of ionospheric disturbances

Participants: Didier Auroux, Sebastian Minjeaud.

In the framework of ANR IODISSEE, in order to identify (and forecast) ionospheric disturbances leading to temporary losses of satellite-to-earth communications (GPS, Galileo), we used Striation software for data assimilation. We obtained the adjoint code thanks to automatic differentiation (Tapenade software from Inria). As the data from Demeter satellite were not available, we extracted synthetic data from a generic model run, and we tried to identified some physical parameters (electronic density, atomic mass, number of particles) of the initial condition from the observations. For a small physical time scale (approximately 1 hour), the identification works very well, and it is possible to retrieve the initial condition from a sparse and noisy observations, allowing us to forecast the evolution of the ionospheric plasma - and then to forecast the disturbances and plasma bubbles that trap GPS and Galileo signals. For longer physical time windows (5 to 10 hours), the identification does not work anymore. We plan to work with real data, if possible, and also with a more complex model (for instance Dynamo software).

CELTIQUE Project-Team

5. New Results

5.1. Browser randomization against web tracking

Participants: Frédéric Besson, Thomas Jensen.

We have investigated different approaches for dynamically tracking information flows in order to improve web browser security. We have identified the problem of stateless web tracking (fingerprinting) and have proposed a novel approach to hybrid information flow monitoring by tracking the knowledge about secret variables using logical formulae. In a follow-up work we investigated how to enforce browser anonymity in the presence of finger-printing web trackers. One way to protect the users' privacy is to make them switch between different machine and browser configurations. We propose a formalisation of this privacy enforcement mechanism. We use information-theoretic channels to model the knowledge of the tracker and the fingerprinting program, and show how to synthesise a randomisation mechanism that defines the distribution of configurations for each user. This mechanism provides a strong guarantee of *privacy* (the probability of identifying the user is bounded by a given threshold) while maximising *usability* (the user switches to other configurations rarely). To find an optimal solution, we express the enforcement problem of randomisation by a linear program. We investigate and compare several approaches to randomisation and find that more efficient privacy enforcement would often provide lower usability. Finally, we relax the requirement of knowing the fingerprinting program in advance, by proposing a randomisation mechanism that guarantees privacy for an arbitrary program.

5.2. Static analysis of functional programs using tree automata and term rewriting

Participants: Thomas Genet, Barbara Kordy, Yann Salmon.

We develop a specific theory and the related tools for analyzing programs whose semantics is defined using term rewriting systems. The analysis principle is based on regular approximations of infinite sets of terms reachable by rewriting. The tools we develop use, so-called, Tree Automata Completion to compute a tree automaton recognizing a superset of all reachable terms. This over-approximation is then used to prove properties on the program by showing that some "bad" terms, encoding dangerous or problematic configurations, are not in the superset and thus not reachable. This is a specific form of, so-called, Regular Tree Model Checking. However, when dealing with infinite-state systems, Regular Tree Model Checking approaches may have some difficulties to represent infinite sets of data. We proposed Lattice Tree Automata, an extended version of tree automata to represent complex data domains and their related operations in an efficient manner. Moreover, we introduce a new completion-based algorithm for computing the possibly infinite set of reachable states in a finite amount of time. This algorithm is independent of the lattice making it possible to seamlessly plug abstract domains into a Regular Tree Model Checking algorithm. These results are part of Valérie Murat's PhD thesis [13]. Now, we aim at applying this technique to the static analysis of programming languages whose semantics is based on terms, like functional programming languages. We already shown that static analysis of first order functional programs can be automated using tree automata completion [28]. Now, one of the objective is to lift those results to the static analysis of higher-order functions. This was precisely the purpose of Yann Salmon's visit to Pr. Luke Ong. Barbara Kordy who joined Celtique in September 2014 is also going to work on this subject.

5.3. Certified JavaScript

Participants: Martin Bodin, Alan Schmitt.

We have completed our first milestone in the development of a certified JavaScript semantics. We have finished a first version of JSCert, a formalization of the current ECMA standard in the Coq proof assistant, and JSRef, a reference interpreter for JavaScript extracted from Coq to OCaml. We have also given a Coq proof that JSRef is correct with respect to JSCert and assessed JSRef using test262, the ECMA conformance test suite. Our methodology ensures that JSCert is a comparatively accurate formulation of the English standard. We have demonstrated that modern techniques of mechanized specification can handle the complexity of JavaScript. This result, obtained in the setting of a collaboration with Philippa Gardner and Sergio Maffeis of Imperial College, and Arthur Charguéraud of Inria Saclay, have been published in the conference Principles of Programming Languages [25].

5.4. SawjaCard: a static analysis tool for certifying Java Card applications

Participants: Frédéric Besson, Thomas Jensen, David Pichardie, Delphine Demange.

We have transfered to the FIME company a static analysis tool for certifying *Java Card* applications, according to security rules defined by the smart card industry. *Java Card* is a dialect of Java designed for programming multi-application smart cards and the tool, called *SawjaCard*, has been specialised for the particular *Java Card* programming patterns. The tool is built around a static analysis engine which uses a combination of numeric and heap analysis. It includes a model of the *Java Card* libraries and the *Java Card* firewall. The tool has been evaluated on a series of industrial applets and is shown to automate a substantial part of the validation process [21].

5.5. Semantics for C programs

Participants: Frédéric Besson, Sandrine Blazy, Pierre Wilke.

Real life C programs are often written using C dialects which, for the ISO C standard, have undefined behaviours. In particular, according to the ISO C standard, reading an uninitialised variable has an undefined behaviour and low-level pointer operations are implementation defined. We propose a formal semantics which gives a well-defined meaning to those behaviours for the C dialect of the CompCert compiler. Our semantics builds upon a novel memory model leveraging a notion of symbolic values. Symbolic values are used by the semantics to delay the evaluation of operations and are normalised lazily to genuine values when needed. We show that the most precise normalisation is computable and that a slightly relaxed normalisation can be efficiently implemented using an SMT solver. The semantics is executable and our experiments show that the enhancements of our semantics are mandatory to give a meaning to low-levels idioms such as those found in the allocation functions of a C standard library [21].

5.6. Fast inference of polynomial invariants

Participants: David Cachera, Thomas Jensen.

We have developed our static analysis techniques for computing polynomial invariants for imperative programs. The analysis is derived from an abstract interpretation of a backwards semantics, and computes preconditions for equalities of the form $g = 0$ to hold at the end of execution. A distinguishing feature of the technique is that it computes polynomial loop invariants without resorting to Gröbner base computations. The analysis uses remainder computations over parameterized polynomials in order to handle conditionals and loops efficiently. The algorithm can analyze and find a large majority of loop invariants reported previously in the literature, and executes significantly faster than implementations using Gröbner bases [15].

5.7. Quantitative analysis of security

Participant: Barbara Kordy.

Graphical models for security is a young but rapidly growing research field. Security models based on graphs combine intuitive, visual representation with rigorous, mathematical foundations. In [30] we address the growing need of performing meaningful probabilistic analysis of security using graphical models. We propose a framework that integrates the modeling technique of attack–defense trees with probabilistic information expressed in terms of Bayesian networks. This allows us to perform probabilistic evaluation of attack–defense scenarios involving dependent actions. To improve the efficiency of our computations, we make use of inference algorithms from Bayesian networks and encoding techniques from constraint reasoning. We discuss the algebraic theory underlying our framework and point out several generalizations which are possible thanks to the use of semiring theory

5.8. Formal Verification of an SSA-Based Middle-End for CompCert

Participants: Delphine Demange, David Pichardie.

CompCert is a formally verified compiler that generates compact and efficient code for a large subset of the C language. However, CompCert foregoes using SSA, an intermediate representation employed by many compilers that enables writing simpler, faster optimizers. In fact, it has remained an open problem to verify formally an SSA-based compiler. We report in [14] on a formally verified, SSA-based middle-end for CompCert. In addition to providing a formally verified SSA-based middle-end, we address two problems raised by Leroy in 2009: giving an intuitive formal semantics to SSA, and leveraging its global properties to reason locally about program optimizations. Joint work with Gilles Barthe.

5.9. A verified information-flow architecture

Participants: Delphine Demange, David Pichardie.

SAFE is a clean-slate design for a highly secure computer system, with pervasive mechanisms for tracking and limiting information flows. At the lowest level, the SAFE hardware supports fine-grained programmable tags, with efficient and flexible propagation and combination of tags as instructions are executed. The operating system virtualizes these generic facilities to present an information-flow abstract machine that allows user programs to label sensitive data with rich confidentiality policies. We present a formal, machine-checked model of the key hardware and software mechanisms used to control information flow in SAFE and an end-to-end proof of noninterference for this model in the Coq proof assistant [17]. This work has been obtained in collaboration with colleagues from University of Pennsylvania, Portland State University, and Harvard University, as part of the CRASH-SAFE project, funded by DARPA.

5.10. Formal Verification of Static Analysis

Participants: Sandrine Blazy, Vincent Laporte, David Pichardie.

Static analysis of binary code is challenging for several reasons. In particular, standard static analysis techniques operate over control flow graphs, which are not available when dealing with self-modifying programs which can modify their own code at runtime. We formalized in the Coq proof assistant some key abstract interpretation techniques that automatically extract memory safety properties and control flow graphs from binary code [22], and operate over a small subset of the x86 assembly. Our analyzer is formally proved correct and has been run on several self-modifying challenges, provided by Cai et al. in their PLDI 2007 paper.

CIDRE Project-Team

6. New Results

6.1. Highlights of the Year

The supervision of distributed system relies heavily on correlation mechanisms that are responsible for collecting alerts coming from sensors and detecting complex scenarios in the flow of alerts. The problem is that it requires to write complex correlation rules. The work we have performed proposes a technique to generate semi-automatically such correlation rules. It describes a process that uses an attack tree and a representation of the system as inputs, and generate a correlation tree that can be translated in an alert correlation description language. This work received the best paper award of SAR-SSI 2014 [50].

One approach to protect the privacy of users in personalized recommendation systems is to publish a sanitized version of the profile of the user by relying a non-interactive mechanism compliant with the concept of differential privacy. In a joint work with Raghavendran Balu and Teddy Furon (LinkMedia Inria team), we have consider two existing schemes offering a differentially private representation of profiles: BLIP (BLoom-and-flIP) and JLT (Johnson-Lindenstrauss Transform). For assessing their security levels, we play the role of an adversary aiming at reconstructing a user profile. To realize this, we design two inference attacks named single and joint decoding. The first inference attack tests the presence of a single item in the profile, and is iterated independently for each possible item of the item set. In contrast, the second inference attack aims at deciding whether a particular subset of items is likely to be in the user profile. This attack is tested on all the possible subsets of items. Our contributions are a theoretical analysis and practical implementations of both attacks tested on datasets composed of real user profiles revealing that joint decoding is the most powerful attack. This also gives useful insights on the setting the differential privacy parameter ϵ. This work has received the best student paper award at the conference ESORICS 2014.
BEST PAPERS AWARDS :
[27] **European Symposium on Research in Computer Security**. R. BALU, T. FURON, S. GAMBS.

6.2. Intrusion Detection

6.2.1. Intrusion detection based on an analysis of information flow control

In 2014, Laurent Georget has started his PhD thesis in the team, working on a subject related to the analysis of information flow control at the kernel level. The goal of his PhD thesis is to propose a formal semantics of the system calls for a real operating systems (namely Linux). This semantics will provide insights about these system calls in terms of information flow. This work will help us to test in a more systematic and efficient way, our reference implementation of a information monitor at the kernel level (Blare).

Blare allows monitoring information flow and identifies the flows that do not conform to a security policy that has been previously defined. Please notice that any explicit flows between OS objects (sockets, files, etc.) are monitored and that in consequence hidden channel attacks cannot be detected by this approach.

We have already developed a dedicated test framework for this software. However, each test written by the developer must be accompanied with the possible results in terms of information flows. The framework simply compares the effective result with the set of expected results. A test passes when the effective result belongs to the set of expected results, and fails otherwise. However, this strategy has turned to be less intuitive than expected. Some system calls must be tested by using several processes operating concurrently. In these cases, the scheduling of processes can produce many different scenarios that will translate quite differently in terms of information flows. To be more confident in our implementation, we really need a stronger and more formal path. The PhD thesis of Laurent Georget is trying to bridge the gap between Blare implementation and the interpretation of the results obtained by running the information flow monitor.

6.2.2. Malware characterization through information flow monitoring

Monitoring information flows consists in observing how pieces of information are disseminated in a given environment. At system level, it consists in intercepting actions performed by an application to deduce how the application disseminates information within the entire operating system. We have propose a new approach to classify and later detect applications infected by malware based on the way they disseminate their own data within an operating system. For this purpose, we first introduce a data-structure named System Flow Graph [thèse Rado to ref.] that offers a compact representation of how pieces of data flow inside a system. A system flow graph describes the external behavior of an application during one execution. Its construction requires no knowledge about the inner working of the application. The graph is built using Blare as an information-flow monitor and more precisely its produced log. We have presented in [25] how these graphs reveal helpful to understand malware behavior and thus why it can help an expert to give a diagnosis in case of intrusion.

6.2.3. Terminating-insensitive non-interference verification based on information flow control

In 2010-2011, we started an informal collaboration with colleagues from CEA LIST laboratory. This collaboration has turned into a reality by the funding of a PhD student (Mounir Assaf). This PhD thesis is about the verification of security properties of programs written in an imperative language with pointer aliasing (a subset of C language) by techniques borrowed from the domain of static analysis. One of the property of interest for the security field is called terminating-insensitive non-interference. Briefly speaking, when verified by a program, this property ensures that the content of any secret variable can not leak into public ones (for any terminating execution). However, this property is too strict in the sense that a large number of programs although perfectly secure are rejected by classical analyzers. Finally in 2014, Mounir Assaf enhanced his previous work on static analysis by introducing a method permitting to quantify information leakage in a C program. This approach requires a theoretical definition of the quantification of information flow leakage and is very promising.

6.2.4. Visualization of security events

The first part of this year was dedicated to tune a working prototype of ELVIS [38] in order to perform field trials with our partner DGA-MI. The prototype was largely well accepted. We were invited by the DGA-MI to present a poster in the Forum DGA Innovation 2014. We will also present ELVIS during the FIC 2014 in Lille on the Pôle Cyber-Défense area.

However, ELVIS also exhibited some limitations of our approach in the way multiple datasets are handled together. We therefore went for a new cycle of research whose objective is to enhance ELVIS in two ways: first to handle multiple datasets at the same time, and second to improve interactions so as to better fit with the processes in forensics. The results of our research lead to CORGI (Combination, Organization and Reconstruction through Graphical Interactions) [39] which was presented at VizSec 2014 (part of Vis 2014). CORGI improves ELVIS by introducing the concepts of *values of interest* that consist in interesting values found by an analyst and that can be used later to search and filter in the other datasets. They are an intuitive and efficient way to link various datasets while the analyst performs its tasks. An early prototype has been developed.

6.2.5. Control flow integrity

In [40] we have studied physical attacks that could disturb the normal execution of an embedded program of a smartcard. Such attacks can be performed using laser beams, electromagnetic glitches and can corrupt the flow of information or change the control flow of the program. We have studied the particular case of the control flow and we have developed software countermeasures that increase the robustness of the control flow. These countermeasures do not require any additional software or hardware external components which is useful for devices like smartcards whose architecture cannot be modified. The developed countermeasures have been validated with the help of the VIS model checker in order to verify that they do not disturb the original execution of the code.

6.2.6. Alert correlation in distributed systems

In large systems, multiple (host and network) Intrusion Detection Systems (IDS) and many sensors are usually deployed. They continuously and independently generate notifications (event's observations, warnings and alerts). To cope with this amount of collected data, alert correlation systems have to be designed. An alert correlation system aims at exploiting the known relationships between some elements that appear in the flow of low level notifications to generate high semantic meta-alerts. The main goal is to reduce the number of alerts returned to the security administrator and to allow a higher level analysis of the situation. However, producing correlation rules is a highly difficult operation, as it requires both the knowledge of an attacker, and the knowledge of the functionalities of all IDSes involved in the detection process. In [50], [47], [36], we focus on the transformation process that allows to translate the description of a complex attack scenario into correlation rules. We show that, once a human expert has provided an action tree derived from an attack tree, a fully automated transformation process can generate exhaustive correlation rules that would be tedious and error prone to enumerate by hand. The transformation relies on a detailed description of various aspects of the real execution environment (topology of the system, deployed services, etc.). Consequently, the generated correlation rules are tightly linked to the characteristics of the monitored information system. The proposed transformation process has been implemented in a prototype that generates correlation rules expressed in an attack description language called Adele.

In the context of the PhD of Mouna Hkimi, we propose a approach to detect intrusions that affect the behavior of distributed applications. To determine whether an observed behavior is normal or not (occurrence of an attack), we rely on a model of normal behavior. This model has been built during an initial training phase. During this preliminary phase, the application is executed several times in a safe environment. The gathered traces (sequences of actions) are used to generate an automaton that characterizes all these acceptable behaviors. To reduce the size of the automaton and to be able to accept more general behaviors that are close to the observed traces, the automaton is transformed. These transformations may lead to introduce unacceptable behaviors. Our current work aims at identifying the possible errors tolerated by the compacted automaton.

6.3. Privacy

6.3.1. Privacy in location-based services

With the advent of GPS-equipped devices, a massive amount of location data is being collected, raising the issue of the privacy risks incurred by the individuals whose movements are recorded. In [17], we focus on a specific inference attack called the de-anonymization attack, by which an adversary tries to infer the identity of a particular individual behind a set of mobility traces. More specifically, we propose an implementation of this attack based on a mobility model called Mobility Markov Chain (MMC). A MMC is built out from the mobility traces observed during the training phase and is used to perform the attack during the testing phase. We design several distance metrics quantifying the closeness between two MMCs and combine these distances to build de-anonymizers that can re-identify users in an anonymized geolocated dataset. Experiments conducted on real datasets demonstrate that the attack is both accurate and resilient to sanitization mechanisms such as downsampling.

One example of a location-based services is dynamic carpooling (also known as instant or ad-hoc ridesharing), which is a service that arranges one-time shared rides on very short notice. This type of carpooling generally makes use of three recent technological advances: (i) navigation devices to determine a route and arrange the shared ride; (ii) smartphones for a traveller to request a ride from wherever she happens to be; and (iii) social networks to establish trust between drivers and passengers. However, the mobiquitous environment in which dynamic carpooling is expected to operate raises several privacy issues. Among all the personal identifiable information, learning the location of an individual is one of the greatest threats against her privacy. For instance, the spatio-temporal data of an individual can be used to infer the location of her home and workplace, to trace her movements and habits, to learn information about her centre of interests or even to detect a change from her usual behavior. Therefore, preserving location privacy is a major issue to be able to leverage the possibilities offered by dynamic carpooling. In a joint work with researchers from LAAS-CNRS

[16], we have propose to follow the privacy-by-design approach by integrating the privacy aspect in the design of dynamic carpooling, henceforth increasing its public (and political) acceptability and trust.

A secure location-based service requires that a mobile user certifies his position before gaining access to a resource. Currently, most of the existing solutions addressing this issue assume a trusted third party that can vouch for the position claimed by a user. However, as computation and communication capacities become ubiquitous with the large scale adoption of smartphones by individuals, these resources can be leverage on to solve this issue in a collaborative and private manner. More precisely together with researchers from LAAS-CNRS, we introduce PROPS, for Privacy-Preserving lOcation Proof System, which allows users to generate proofs of location in a private and distributed way using neighboring nodes as witnesses [35]. PROPS provides security properties such as unforgeability and non-transferability of the proofs, as well as resistance to classical localization attacks.

One of the fundamental building block to construct a location proof system such as PROPS is a distance-bounding protocol. More precisely, in distance-bounding authentication protocols a verifier assesses that a prover is (1) legitimate and (2) in the verifier's proximity. Proximity checking is done by running time-critical exchanges between both parties. This enables the verifier to detect relay attacks (also called mafia fraud). While most distance-bounding protocols offer resistance to mafia, distance, and impersonation attacks, only few protect the privacy of the authenticating prover. One exception is the protocol due to Hermans, Peeters, and Onete, which offers prover untraceability with respect to a Man-in-the-Middle adversary. However in this protocol as well as in all other distance-bounding protocols, any legitimate verifier can identify, and thus track, the prover. In order to counter the threats of possible corruption or data leakage from verifiers, together with Jean-Marc Robert (ETS, Montréal) we propose a distance-bounding protocol providing strong prover privacy with respect to the verifier and deniability with respect to a centralized back-end server managing prover creation and revocation [33]. In particular, we first formalize the notion of prover anonymity, which guarantees that even verifiers cannot trace provers, and deniability, which allows provers to deny that they were authenticated by a verifier. Finally, we prove that our protocol achieves these strong guarantees.

A particular class of relay attacks against distance-bounding protocols is called terrorist fraud in which a distant malicious prover colludes with an attacker located in a verier's proximity when authenticating. Existing distance-bounding protocols resisting such attacks are designed to be lightweight and thus symmetric, relying on a secret shared by the prover and the verifier. Recently, several asymmetric distance-bounding protocols were proposed by Gambs, Onete and Robert as well as by Hermans, Peter and Onete, but they fail to thwart terrorist fraud. One earlier asymmetric protocol aiming to be terrorist-fraud resistant is the DBPK-Log protocol due to Bussard and Bagga, which was unfortunately recently proven to achieve neither distance- nor terrorist-fraud resistance. In this work, we build on some ideas of the DBPK-Log scheme and propose a novel distance-bounding protocol resistant to terrorist fraud that does not require the pre-existence of a shared secret between the prover and the verifier [32]. Our construction, denoted as VSSDB (for Verifiable Secret Sharing and Distance-Bounding Protocol) relies on a veriable secret sharing scheme and on the concept of modes, which we introduce as a novel element to complement fast-round challenges in order to improve security. We prove that VSSDB achieves terrorist-fraud resistance in a relaxed security model called KeyTF-security, which we also present in this paper.

6.3.2. Equity in privacy-enhanced social networks

In [46], we have examined a novel issue in the field of policy conflict resolution, and applied it to privacy policy management in distributed social networking systems. We accepted as a starting point that in a privacy-enhanced social network, when a user publishes a document (e.g., a picture), any user referenced in this document (e.g., people tagged in pictures) should be entitled to issue a privacy policy over this document. In this case, when a given user tries to access a given document, multiple users may issue multiple access control decisions (or rulings), possibly resulting in a normative conflict. Quite a number of strategies are available for the resolution of such conflicts, the most common one being the "deny strategy", allowing any ruling denying access to the resource to take precedence over others. This is usually considered a "secure" way of dealing with access control. However, with this strategy as with many others, it is possible for a user to design her policy in a way that systematically prevents other users from interacting in a normal way, while allowing herself to potentially benefit from other people's more flexible policies. This may leads to unfair situations, in which some users take advantage of the systems while others' experience is damaged. This is particularly an issue in social networking applications, in which information sharing is a core feature and access restrictions, while necessary to protect intimacy, can sometimes be considered aggressive.

To address this particular trade-off between privacy and usability, we have introduced the notion of equity in such scenarios, a situation being equitable when all involved users have seen their policy enforced or violated in the same proportion over past interactions. We have designed a conflict resolution algorithm aimed at improving this equity in our social networking scenario, and evaluated its impact by measuring Gini coefficients (an indicator commonly used by economists to measure the distribution of wealth in a population) over the distribution of enforcement proportions in the population of users. With respect to this criterion, it actually proved more efficient than other strategies. Following these positive results, we have recently taken steps towards a formalization and generalization of this intuitive concept of equity and the design of systematic tools to evaluate and compare the impact of any conflict resolution strategy over various possible flavors of the notion.

6.3.3. Private mobile services

The development of NFC-enabled smartphones has paved the way to new applications such as mobile payment (m-payment) and mobile ticketing (m-ticketing). However, often the privacy of users of such services is either not taken into account or based on simple pseudonyms, which does not offer strong privacy properties such as the unlinkability of transactions and minimal information leakage. In [48], [15], we introduce a lightweight privacy-preserving contactless transport service that uses the SIM card as a secure element. Our implementation of this service uses a group signature protocol in which costly cryptographic operations are delegated to the mobile phone. We have also conducted an interdisciplinary study with researchers from social sciences to analyze the media coverage in the modern public space on the topic of privacy with respect to mobile technologies [29]. Despite the difficulties highlighted by these studies, we argue that research efforts should support the emergence of mobile services that respect users' privacy as well as the development of a digital culture of privacy.

6.3.4. Architectures for privacy

In the current architecture of the Internet, there is a strong asymmetry in terms of power between the entities that gather and process personal data (e.g., major Internet companies, telecom operators, cloud providers, ...) and the individuals from which this personal data is issued. In particular, individuals have no choice but to blindly trust that these entities will respect their privacy and protect their personal data. In a position paper [34] in a collaboration with researchers from the Université de Montréal and Aarhus University, we propose an utopian crypto-democracy model based on existing scientific achievements from the field of cryptography. More precisely, our main objective is to show that cryptographic primitives, including in particular secure multiparty computation, offer a practical solution to protect privacy while minimizing the trust assumptions. In the crypto-democracy envisioned, individuals do not have to trust a single physical entity with their personal data but rather their data is distributed among several institutions. Together these institutions form a virtual entity called the Trustworthy that is responsible for the storage of this data but which can also compute on it (provided first that all the institutions agree on this). Finally, we also propose a realistic proof-of-concept of the Trustworthy, in which the roles of institutions are played by universities. This proof-of-concept would have an important impact in demonstrating the possibilities offered by the crypto-democracy paradigm.

Active fingerprinting schemes were originally invented to deter malicious users from illegally releasing an item, such as a movie or an image. To achieve this, each time an item is released, a different fingerprint is embedded in it. If the fingerprint is created from an anti-collusion code, the fingerprinting scheme can trace colluding buyers who forge fake copies of the item using their own legitimate copies. Charpentier, Fontaine, Furon and Cox were the first to propose an asymmetric fingerprinting scheme based on Tardos codes, the most efficient anti-collusion codes known to this day. However, their work focuses on security but does not preserve the privacy of buyers. To address this issue, we introduce the first privacy-preserving asymmetric fingerprinting protocol based on Tardos codes [30]. This protocol is optimal with respect traitor tracing. We also formally define the properties of correctness, anti-framing, traitor tracing, as well as buyer- and item-unlinkability. Finally, we prove that our protocol achieves these properties and give exact bounds for each of them.

6.3.5. Privacy and web services

We have proposed [61] a new model of security policy based for a first part on our previous works in information flow policy and for a second part on a model of Myers and Liskov. This new model of information flow serves web services security and allows a user to precisely define where its own sensitive pieces of data are allowed to flow through the definition of an information flow policy. A novel feature of such policy is that they can be dynamically updated, which is fundamental in the context of web services that allow the dynamic discovery of services. We have also presented an implementation of this model in a web services orchestration in BPEL (Business Process Execution Language).

6.3.6. Privacy-preserving ad-hoc routing

Last year, we have proposed NoName, a privacy-preserving ad-hoc routing protocol. Based on trapdoor, virtual switching and partially disjoint multipaths using Bloom filter, NoName ensures the anonymity of the source, of the destination and of intermediate nodes. It also ensures unlinkability between source and message and between destination and message. Since then, we have demonstrated that colluding attackers analyzing Bloom filters can locate the origin node of routes requests messages. Thus, Noname, like ARMR, another privacy-preserving ad-hoc routing protocol using also Bloom filter, do not prevent the localization of the source. We have developed a cryptographic primitive called fuzzy cryptographic Bloom filter that offers the same functions as Bloom filters (in our case, preventing routing loops) while preventing localization of the source of route request messages.

6.4. Trust

Digital reputation mechanisms have indeed emerged as a promising approach to cope with the specificities of large scale and dynamic systems. Similarly to real world reputation, a digital reputation mechanism expresses a collective opinion about a target user based on aggregated feedback about his past behavior. The resulting reputation score is usually a mathematical object (*e.g.* a number or a percentage). It is used to help entities in deciding whether an interaction with a target user should be considered. Digital reputation mechanisms are thus a powerful tool to incite users to behave trustworthily. Indeed, a user who behaves correctly improves his reputation score, encouraging more users to interact with him. In contrast, misbehaving users have lower reputation scores, which makes it harder for them to interact with other users. To be useful, a reputation mechanism must itself be accurate against adversarial behaviors. Indeed, a user may attack the mechanism to increase his own reputation score or to reduce the reputation of a competitor. A user may also free-ride the mechanism and estimate the reputation of other users without providing his own feedback. From what has been said, it should be clear that reputation is beneficial in order to reduce the potential risk of communicating with almost or completely unknown entities. Unfortunately, the user privacy may easily be jeopardized by reputation mechanisms, which is clearly a strong argument to compromise the use of such a mechanism. Indeed, by collecting and aggregating user feedback, or by simply interacting with someone, reputation systems can be easily manipulated in order to deduce user profiles. Thus preserving user privacy while computing robust reputation is a real and important issue that we address in our work [51]. Specifically, our proposal aims at enhancing signatures of reputation mechanism proposed by Bethencourt and his colleagues in 2010 by handling negative votes. Taking into account negative votes implies major modifications with respect to the implementation of the mechanism. Specifically, in the mechanism of Bethencourt and co-authors, service providers locally store votes cast at the end of their interaction with their clients, and compute their reputation score by aggregating the received votes. In particular, they can keep only a subset of them, which clearly makes negative votes useless. We propose to improve upon this solution by guaranteeing that negative votes are taken into account. This is achieved by making both reputation scores and votes of service providers publicly available in order to prevent anyone from modifying or hiding them. Our proposition accomplishes this without jeopardizing the privacy of clients.

6.5. Other topics related to security and distributed computing

6.5.1. Network monitoring and fault detection

Monitoring a system consists in collecting and analyzing relevant information provided by the monitored devices, so as to be continuously aware of the system state (situational awareness). However, the ever growing complexity and scale of systems makes both real time monitoring and fault detection a quite tedious task. Thus the usually adopted option is to focus solely on a subset of information states, so as to provide coarse-grained indicators. As a consequence, detecting isolated failures or anomalies is a quite challenging issue. We propose in [23], [42] to address this issue by pushing the monitoring task at the edge of the network. We present a peer-to-peer based architecture, which enables nodes to adaptively and efficiently self-organize according to their "health" indicators. By exploiting both temporal and spatial correlations that exist between a device and its vicinity, our approach guarantees that only isolated anomalies (an anomaly is isolated if it impacts solely a monitored device) are reported on the fly to the network operator. We show that the end-to-end detection process, *i.e.*, from the local detection to the management operator reporting, requires a logarithmic number of messages in the size of the network.

6.5.2. Secure data deduplication scheme

Data grows at the impressive rate of 50% per year, and 75% of the digital world is a copy [0]. Although keeping multiple copies of data is necessary to guarantee their availability and long term durability, in many situations the amount of data redundancy is immoderate. By keeping a single copy of repeated data, data deduplication is considered as one of the most promising solutions to reduce the storage costs, and improve users experience by saving network bandwidth and reducing backup time. However, this solution must now solve many security issues to be completely satisfying. In this paper we target the attacks from malicious clients that are based on the manipulation of data identifiers and those based on backup time and network traffic observation. In [43], we have presented a deduplication scheme mixing an intra-and an inter-user deduplication in order to build a storage system that is secure against the aforementioned type of attacks by controlling the correspondence between files and their identifiers, and making the inter-user deduplication unnoticeable to clients using deduplication proxies. Our method provides global storage space savings, per-client bandwidth network savings between clients and deduplication proxies, and global network bandwidth savings between deduplication proxies and the storage server. The evaluation of our solution compared to a classic system shows that the overhead introduced by our scheme is mostly due to data encryption which is necessary to ensure confidentiality. This work relies on Mistore [44], [45], a distributed storage system aiming at guaranteeing data availability, durability, low access latency by leveraging the Digital Subscriber Line infrastructure of an ISP. Mistore uses the available storage resources of a large number of home gateways and points of presence for content storage and caching facilities reducing the role of the data center to a load balancer. Mistore also targets data consistency by providing multiple types of consistency criteria on content and a versioning system allowing users to get access to any prior versions of their contents.

6.5.3. Metrics estimation on very large data streams

In [12], we consider the setting of large scale distributed systems, in which each node needs to quickly process a huge amount of data received in the form of a stream that may have been tampered with by an adversary (*i.e.*, data items ordering can be manipulated by an oblivious adversary). In this situation, a fundamental problem is how to detect and quantify the amount of work performed by the adversary. To address this issue, we propose AnKLe (for Attack-tolerant eNhanced Kullback- Leibler divergence Estimator), a novel algorithm for estimating the KL divergence of an observed stream compared to the expected one. AnKLe combines sampling techniques and information-theoretic methods. It is very efficient, both in terms of space and time complexities, and requires only a single pass over the data stream. Experimental results show that the estimation provided by AnKLe remains accurate even for different adversarial settings for which the quality of other methods dramatically decreases. Considering n as the number of distinct data items in a stream, we show that AnKLe is an (ε, δ)-approximation algorithm with a space complexity sublinear in the size of the domain value from which data items are drawn and the maximal stream length.

[0] The digital universe decade. Are you ready? John Gantz and David Reinsel, IDC information, may 2010.

We go a step further by proposing in [22] a metric, called codeviation, that allows to evaluate the correlation between distributed streams. This metric is inspired from classical metric in statistics and probability theory, and as such allows us to understand how observed quantities change together, and in which proportion. We then propose to estimate the codeviation in the data stream model. In this model, functions are estimated on a huge sequence of data items, in an online fashion, and with a very small amount of memory with respect to both the size of the input stream and the values domain from which data items are drawn. We give upper and lower bounds on the quality of the codeviation, and provide both local and distributed algorithms that additively approximates the codeviation among n data streams by using a sublinear number of bits of space in the size of the domain value from which data items are drawn and the maximal stream length. To the best of our knowledge, such a metric has never been proposed so far.

6.5.4. Robustness analysis of large scale distributed systems

In the continuation of [59] which proposed an in-depth study of the dynamicity and robustness properties of large-scale distributed systems, we analyze in [13], the behavior of a stochastic system composed of several identically distributed, but non independent, discrete-time absorbing Markov chains competing at each instant for a transition. The competition consists in determining at each instant, using a given probability distribution, the only Markov chain allowed to make a transition. We analyze the first time at which one of the Markov chains reaches its absorbing state. When the number of Markov chains goes to infinity, we analyze the asymptotic behavior of the system for an arbitrary probability mass function governing the competition. We give conditions for the existence of the asymptotic distribution and we show how these results apply to cluster-based distributed systems when the competition between the Markov chains is handled by using a geometric distribution.

6.5.5. Detection of distributed denial-of-service attacks

A Denial-of-Service (DoS) attack tries to progressively take down an Internet resource by flooding this resource with more requests than it is capable to handle. A Distributed Denial-of-Service (DDoS) attack is a DoS attack triggered by thousands of machines that have been infected by a malicious software, with as immediate consequence the total shut down of targeted web resources (*e.g.*, e-commerce websites). A solution to detect and to mitigate DDoS attacks it to monitor network traffic at routers and to look for highly frequent signatures that might suggest ongoing attacks. A recent strategy followed by the attackers is to hide their massive flow of requests over a multitude of routes, so that locally, these flows do not appear as frequent, while globally they represent a significant portion of the network traffic. The term "iceberg" has been recently introduced to describe such an attack as only a very small part of the iceberg can be observed from each single router. The approach adopted to defend against such new attacks is to rely on multiple routers that locally monitor their network traffic, and upon detection of potential icebergs, inform a monitoring server that aggregates all the monitored information to accurately detect icebergs. To prevent the server from being overloaded by all the monitored information, routers continuously keep track of the c (among n) most recent high flows (modeled as items) prior to sending them to the server, and throw away all the items that appear with a small probability p_i, and such that the sum of these small probabilities is modeled by probability p_0. Parameter c is dimensioned so that the frequency at which all the routers send their c last frequent items is low enough to enable the server to aggregate all of them and to trigger a DDoS alarm when needed. This amounts to compute the time needed to collect c distinct items among n frequent ones. A thorough analysis of the time needed to collect c distinct items appears in [53].

6.5.6. Randomized message-passing test-and-set

In [56], we have presented a solution to the well-known Test&Set operation in an asynchronous system prone to process crashes. Test&Set is a synchronization operation that, when invoked by a set of processes, returns yes to a unique process and returns no to all the others. Recently many advances in implementing Test&Set objects have been achieved, however all of them target the shared memory model. In this paper we propose an implementation of a Test&Set object in the message passing model. This implementation can be invoked by any number $p < n$ of processes in which n is the total number of processes in the system. It has an expected individual step complexity in $O(\log p)$ against an oblivious adversary, and an expected individual message complexity in $O(n)$. The proposed Test&Set object is built atop a new basic building block, called selector, that allows to select a winning group among two groups of processes. We propose a message-passing implementation of the selector whose step complexity is constant. We are not aware of any other implementation of the Test&Set operation in the message passing model.

6.5.7. Agreement problems in unreliable systems

In [18], we consider the problem of approximate consensus in mobile ad-hoc networks in the presence of Byzantine nodes. Each node begins to participate by providing a real number called its initial value. Eventually all correct nodes must obtain final values that are different from each other within a maximum value previously defined (convergence property) and must be in the range of initial values proposed by the correct nodes (validity property). Due to nodes' mobility, the topology is dynamic and unpredictable. We propose an approximate Byzantine consensus protocol which is based on the linear iteration method. Each node repeatedly executes rounds. During a round, a node moves to a new location, broadcasts its current value, gathers values from its neighbors, and possibly updates its value. In our protocol, nodes are allowed to collect information during several consecutive rounds: thus moving gives them the opportunity to gather progressively enough values. An integer parameter R_c is used to define the maximal number of rounds during which values can be gathered and stored while waiting to be used. A novel sufficient and necessary condition guarantees the final convergence of the consensus protocol. At each stage of the computation, a single correct node is concerned by the requirement expressed by this new condition (the condition is not universal as it is the case in all previous related works). Moreover the condition considers both the topology and the values proposed by correct nodes. If less than one third of the nodes are faulty, the condition can be satisfied. We are working on mobility scenarios (random trajectories, predefined trajectories, meeting points) to assert that the condition can be satisfied for reasonable values of R_c. In [41], we extend the above protocol to solve the problem of clock synchronization in mobile ad-hoc networks.

In [20], we investigate the use of agreement protocols to develop transactional mobile agents. Mobile devices are now equipped with multiple sensors and networking capabilities. They can gather information about their surrounding environment and interact both with nearby nodes, using a dynamic and self-configurable ad-hoc network, and with distant nodes via the Internet. While the concept of mobile agent is appropriate to explore the ad-hoc network and autonomously discover service providers, it is not suitable for the implementation of strong distributed synchronization mechanisms. Moreover, the termination of a task assigned to an agent may be compromised if the persistence of the agent itself is not ensured. In the case of a transactional mobile agent, we identify two services, Availability of the Sources and Atomic Commit, that can be supplied by more powerful entities located in a cloud. We propose a solution in which these two services are provided in a reliable and homogeneous way. To guarantee reliability, the proposed solution relies on a single agreement protocol that orders continuously all the new actions whatever the related transaction and service.

CLASSIC Project-Team

4. New Results

4.1. Corpus linguistics and Markov substitute processes

Thomas Mainguy and Olivier Catoni studied a new statistical model for natural language modeling, called Markov substitute processes. This model is based on a set of conditional independence properties that are more general than the Markov field assumption. It has connections with context free grammars and forms a collection of exponential families having for this reason nice estimation properties.

4.2. Kernel Principal Component Analysis and spectral clustering

Ilaria Giulini and Olivier Catoni continued their study of dimension free bounds for the estimation of the Gram matrix and more generally for the estimation of the expectation of a random symmetric matrix from an i.i.d. sample. This study, using PAC-Bayes bounds, both leads to new robust estimators with applications to Principal Component Analysis in high of even infinite dimension, and new bounds for the usual empirical Gram matrix estimate. Getting dimension free bounds is important to get new results on Kernel PCA. Applications were also studied to density estimation and to spectral clustering.

CLIME Project-Team

6. New Results

6.1. Highlights of the Year

BEST PAPER AWARD :

[20] **VISAPP - International Conference on Computer Vision Theory and Applications**. D. BÉRÉZIAT, I. HERLIN.

6.2. State estimation: analysis and forecast

One major objective of Clime is the conception of new methods of data assimilation in geophysical sciences. Clime is active on several challenging aspects: non-Gaussian assumptions, multiscale assimilation, minimax filtering, etc.

6.2.1. An iterative ensemble Kalman smoother

Participants: Marc Bocquet, Pavel Sakov [BOM, Australia].

The iterative ensemble Kalman filter (IEnKF) was proposed for improving the performance of the ensemble Kalman filter on strongly nonlinear geophysical models. IEnKF can be used as a lag-one smoother and extended to a fixed-lag smoother: the iterative ensemble Kalman smoother (IEnKS). IEnKS is an ensemble variational method. It does not require the use of the tangent of the evolution and observation models, nor the adjoint of these models: the required sensitivities (gradient and Hessian) are computed from the ensemble. Looking for the optimal performance, we consider a quasi-static algorithm, out of the many possible extensions. IEnKS was explored on the Lorenz'95 model and on a 2D turbulence model. As a logical extension of IEnKF, IEnKS significantly outperforms the standard Kalman filters and smoothers in strongly nonlinear regimes. In mildly nonlinear regimes (typically synoptic scale meteorology), its filtering performance is marginally but clearly better than the standard ensemble Kalman filter, and it keeps improving as the length of the temporal data assimilation window is increased. For long windows, its smoothing performance very significantly outranks the standard smoothers, which is believed to stem from the variational but flow-dependent nature of the algorithm. For very long windows, the use of a multiple data assimilation variant of the scheme, where observations are assimilated several times, is advocated. This paves the way for finer re-analysis freed from the static prior assumption of 4D-Var, but also partially freed from the Gaussian assumptions that usually impede standard ensemble Kalman filtering and smoothing.

6.2.2. Modeling and assimilation of lidar signals

Participants: Yiguo Wang [CEREA], Karine Sartelet [CEREA], Marc Bocquet, Patrick Chazette [LSCE, France].

In this study, we investigate the ability of the chemistry transport model (CTM) Polair3D of the air quality platform Polyphemus to simulate lidar backscattered profiles from model aerosol concentration outputs. This investigation is an important pre-processing stage of data assimilation (validation of the observation operator). To do so, simulated lidar signals are compared to hourly lidar observations performed during the MEGAPOLI (Megacities: Emissions, urban, regional and Global Atmospheric POLlution and climate effects, and Integrated tools for assessment and mitigation) summer experiment in July 2009, when a ground-based mobile lidar was deployed around Paris on-board a van. The comparison is performed for six days (1, 4, 16, 21, 26 and 29 July 2009), corresponding to different levels of pollution and different atmospheric conditions. Overall, Polyphemus reproduces well the vertical distribution of lidar signals and their temporal variability, especially for 1, 16, 26 and 29 July 2009. Discrepancies on 4 and 21 July 2009 are due to high-altitude aerosol layers, which are not well modeled. In the second part of this study, two new algorithms for assimilating lidar observations based on the optimal interpolation method are presented. One algorithm analyses PM_{10} (particulate matter with diameter less than 10 μm) concentrations. Another analyses $PM_{2.5}$ (particulate matter with diameter less than 2.5 μm) and $PM_{2.5-10}$ (particulate matter with a diameter higher than 2.5 μm and lower than 10 μm) concentrations separately. The aerosol simulations without and with lidar Data Assimilation (DA) are evaluated using the Airparif (a regional operational network in charge of air quality survey around the Paris area) database to demonstrate the feasibility and usefulness of assimilating lidar profiles for aerosol forecasts. The evaluation shows that lidar DA is more efficient at correcting PM_{10} than $PM_{2.5}$, probably because $PM_{2.5}$ is better modeled than PM_{10}. Furthermore, the algorithm which analyzes both $PM_{2.5}$ and $PM_{2.5-10}$ provides the best scores for PM_{10}. The averaged root-mean-square error (RMSE) of PM_{10} is 11.63 $\mu g\, m^{-3}$ with DA ($PM_{2.5}$ and $PM_{2.5-10}$), compared to 13.69 $\mu g\, m^{-3}$ with DA (PM_{10}) and 17.74 $\mu g\, m^{-3}$ without DA on 1 July 2009. The averaged RMSE of PM_{10} is 4.73 $\mu g\, m^{-3}$ with DA ($PM_{2.5}$ and $PM_{2.5-10}$), against 6.08 $\mu g\, m^{-3}$ with DA (PM_{10}) and 6.67 $\mu g\, m^{-3}$ without DA on 26 July 2009.

6.2.3. Assimilation of lidar signals: application to aerosol forecasting

Participants: Yiguo Wang [CEREA], Karine Sartelet [CEREA], Marc Bocquet, Patrick Chazette [LSCE].

This study represents a new application of assimilating lidar signals to aerosol forecasting. It aims at investigating the impact of a ground-based lidar network on the analysis and short-term forecasts of aerosols through a case study in the Mediterranean basin. To do so, we employ a Data Assimilation (DA) algorithm based on the optimal interpolation method developed in the Polair3D chemistry transport model (CTM) of the Polyphemus air quality modeling platform. We assimilate hourly averaged normalized range-corrected lidar signals retrieved from a 72 h period of intensive and continuous measurements performed in July 2012 by ground-based lidar systems of the European Aerosol Research Lidar Network (EARLINET). Particles with an aerodynamic diameter lower than 2.5 μm ($PM_{2.5}$) and those with an aerodynamic diameter higher than 2.5 μm but lower than 10 ($PM_{10-2.5}$) are analyzed separately using the lidar observations at each DA step. First, we study the spatial and temporal influences of the assimilation of lidar signals on aerosol forecasting. We conduct sensitivity studies on algorithmic parameters, e.g. the horizontal correlation length (L_h) used in the background error covariance matrix (50 km, 100 km or 200 km), the altitudes at which DA is performed (0.75–3.5 km, 1.0–3.5 km or 1.5–3.5 km) and the assimilation period length (12 h or 24 h). We find that DA with $L_\mathrm{h} = 100$ km and assimilation from 1.0 to 3.5 km during a 12 h assimilation period length leads to the best scores for PM_{10} and $PM_{2.5}$ during the forecast period with reference to available measurements from surface networks. Secondly, the aerosol simulation results without and with lidar DA using the optimal parameters (L_h = 100 km, an assimilation altitude range from 1.0 to 3.5 km and a 12 h DA period) are evaluated using the level 2.0 (cloud-screened and quality-assured) aerosol optical depth data from AERONET, and mass concentration measurements (PM_{10} or $PM_{2.5}$) from the French air quality (BDQA) network and the EMEP-Spain/Portugal network. The results show that the simulation with DA leads to better scores than the one without DA for $PM_{2.5}$, PM_{10} and aerosol optical depth. Additionally, the comparison of model results to evaluation data indicates that the temporal impact of assimilating lidar signals is longer than 36 h after the assimilation period.

Fig. 2 shows the performance of assimilating real lidar data over the Mediterranean sea with a view to forecast particulate matter over France.

6.2.4. Local ensemble transform Kalman filter for adaptive optics on extremely large telescopes

Participants: Morgan Gray [LAM, France], Cyril Petit [ONERA, France], Sergei Rodionov [LAM, France], Marc Bocquet, Laurent Bertino [NERSC, Norway], Marc Ferrari [LAM, France], Thierry Fusco [LAM and ONERA, France].

We proposed a new algorithm for an adaptive optics system control law, based on the Linear Quadratic Gaussian approach and a Kalman Filter adaptation with localizations. It allows to handle non-stationary behaviors, to obtain performance close to the optimality defined with the residual phase variance minimization criterion, and to reduce the computational burden with an intrinsically parallel implementation on the Extremely Large Telescopes.

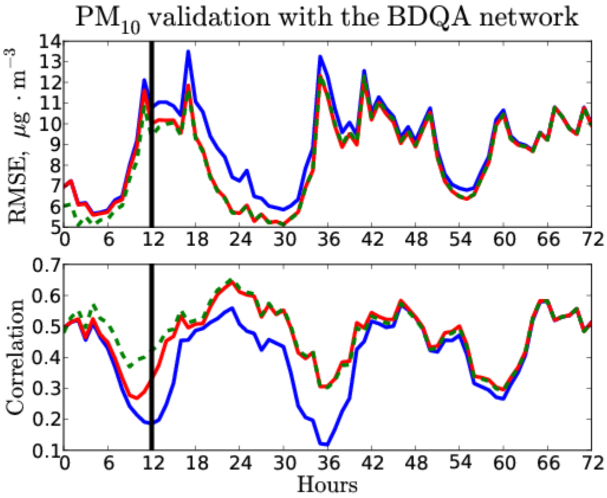


Figure 2. Validation of forecasts of particulate matter PM_{10} using ground stations over France when lidar data have been assimilated over the Mediterranean sea. These forecasts (red line: 12-hour assimilation period and dashed green line: 24-hour assimilation period) are compared to a free run (blue line).

6.3. Inverse modeling

Research on inverse modeling techniques is a major component of Clime, with a focus, in 2014, on hyperparameter estimation when the statistics are non-Gaussian.

6.3.1. Estimation of the caesium-137 source term from the Fukushima Daiichi plant

Participants: Victor Winiarek, Marc Bocquet, Nora Duhanyan [CEREA], Yelva Roustan [CEREA], Olivier Saunier [IRSN], Anne Mathieu [IRSN].

To estimate the amount of radionuclides and the temporal profile of the source term released in the atmosphere during the accident of the Fukushima Daiichi nuclear power plant in March 2011, inverse modeling techniques have been used and have proven their ability in this context. In a previous study, the lower bounds of the caesium-137 and iodine-131 source terms were estimated with such techniques, using activity concentration observations. The importance of an objective assessment of prior errors (the observation errors and the background errors) was emphasized for a reliable inversion. In such critical context where the meteorological conditions can make the source term partly unobservable and where only a few observations are available, such prior estimation techniques are mandatory, the retrieved source term being very sensitive to this estimation.

We propose to extend the use of these techniques to the estimation of prior errors when assimilating observations from several data sets. The aim is to compute an estimate of the caesium-137 source term jointly using all available data about this radionuclide, such as activity concentrations in the air, but also daily fallout measurements and total cumulated fallout measurements. It is crucial to properly and simultaneously estimate the background errors and the prior errors relative to each data set. A proper estimation of prior errors is also a necessary condition to reliably estimate the a posteriori uncertainty of the estimated source term. Using such techniques, we retrieve a total released quantity of caesium-137 in the interval $11.6 - 19.3$ PBq with an estimated standard deviation range of $15 - 20\%$ depending on the method and the data sets. The “blind” time intervals of the source term have also been strongly mitigated compared to the first estimations with only activity concentration data.

6.4. Image assimilation

Sequences of images, such as satellite acquisitions, display structures evolving in time. This information is recognized of major interest by forecasters (meteorologists, oceanographers, etc.) in order to improve the information provided by numerical models. However, the satellite images are mostly assimilated in geophysical models on a point-wise basis, discarding the space-time coherence visualized by the evolution of structures such as clouds. Assimilating in an optimal way image data is of major interest and this issue should be considered in two ways:

- from the model's viewpoint, the location of structures on the observations is used to control the state vector.
- from the image's viewpoint, a model of the dynamics and structures is built from the observations.

6.4.1. Model error and motion estimation

Participants: Dominique Béréziat [UPMC], Isabelle Herlin.

Data assimilation technics are used to retrieve motion from image sequences. These methods require a model of the underlying dynamics, displayed by the evolution of image data. In order to quantify the approximation linked to the chosen dynamic model, an error term is included in the evolution equation of motion and a weak formulation of 4D-Var data assimilation is designed. The cost function to be minimized depends simultaneously on the initial motion field, at the beginning of the studied temporal window, and on the error value at each time step. The result allows to assess the model error and analyze its impact on motion estimation. The approach is used to estimate geophysical forces (gravity, Coriolis, diffusion) from images in order to better assess the surface dynamics [20] and forecast the displacement of structures like oilspill.

6.4.2. Tracking of structures from an image sequence

Participants: Yann Lepoittevin, Isabelle Herlin, Dominique Béréziat [UPMC].

The research concerns an approach to estimate velocity on an image sequence and simultaneously segment and track a given structure. It relies on the underlying dynamics' equations of the studied physical system. A data assimilation method is designed to solve evolution equations of image brightness, those of motion's dynamics, and those of the distance map modeling the tracked structures. The method is applied on meteorological satellite data, in order to track tropical clouds on image sequences and estimate their motion, as seen on Fig. 3 .

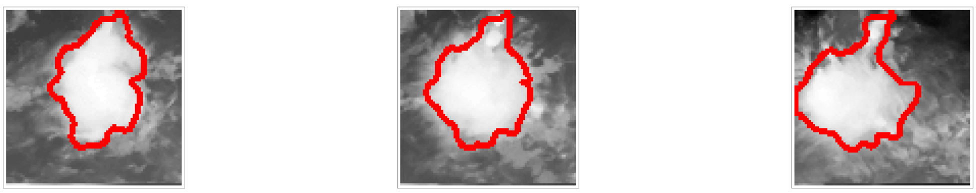


Figure 3. Tracking a tropical cloud. Frames 3, 9, 18 of the sequence.

Quantification is obtained on synthetic experiments by comparing trajectories of characteristic points. The respective position of these points on the last image of the sequence for different methods may be compared to that obtained with ground truth as seen on Fig. 4 .

Data assimilation is performed either with a 4D-Var variational approach or with a Kalman ensemble method [22]. In the last case, the initial ensemble is obtained from a set of optical flow methods of the literature with various parameters values.

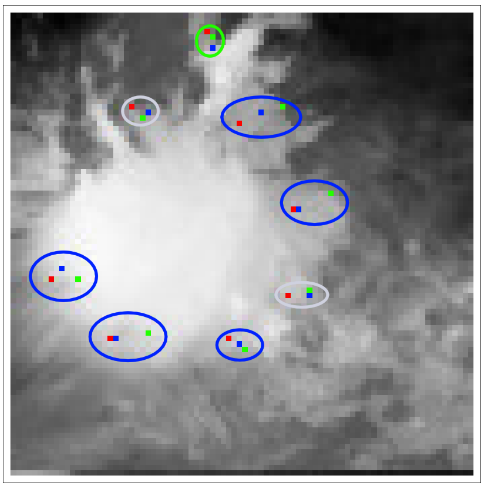


Figure 4. Red point: ground truth. Blue point: our method. Green point: Sun's optical flow. Blue ellipse: our method is the best. Green ellipse: Sun's result is the best. Grey ellipse : results are equivalent.

6.4.3. Motion estimation from images with a waveforms reduced model

Participants: Etienne Huot, Isabelle Herlin, Giuseppe Papari [CFLIR, Belgium].

Dimension reduction is applied to a model of image evolution, composed of transport of velocity and image brightness. Waveform bases are obtained on the image domain for subspaces of images and motion fields, as eigenvectors of previously defined quadratic functions. Image assimilation with the reduced model allows to estimate velocity fields satisfying the space-time properties chosen defined by the user for designing the quadratic function. This approach allows complex geographical domains and suppresses the difficulty of boundary conditions on such domains: these boundary conditions are automatically applied on the bases elements. Motion estimation is then obtained with a reduced model whose state vector is composed of a few components for motion and images. This has to be compared with the initial motion estimation problem that involves a state vector that has a size proportional to the image domain. Current research concern the definition of new quadratic functions from image properties.

6.4.4. Applying POD on a model output dabase for defining a reduced motion model

Participants: Etienne Huot, Isabelle Herlin.

Dimension reduction may also be studied by determining a small size reduced basis obtained by Proper Orthogonal Decomposition (POD) of a motion fields database. This database is constructed for characterizing accurately the surface circulation of the studied area, so that linear combinations of the basis elements obtained by POD accurately describe the motion function observed on satellite image sequences. The database includes the geostrophic motion fields obtained from Sea Level Anomaly reanalysis maps that are available from the MyOcean European project website (http://www.myocean.eu/). Fig. 5 displays such SLA maps and the associated motion fields.

Image assimilation with the POD reduced model allows estimating motion as displayed on Fig. 6 .

6.4.5. Rain nowcasting from radar image acquisitions

Participants: Yann Lepoittevin, Isabelle Herlin.

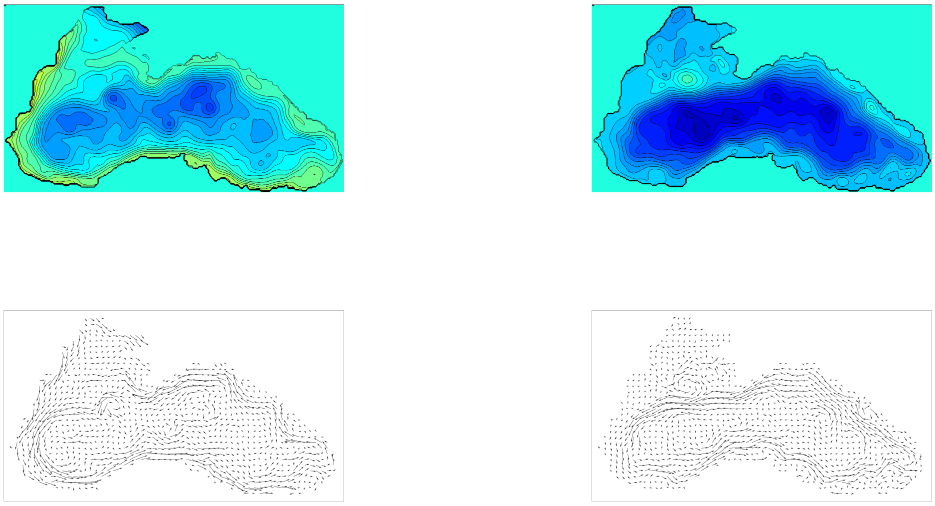


Figure 5. Top: reanalysis of SLA. Bottom: geostrophic motion.

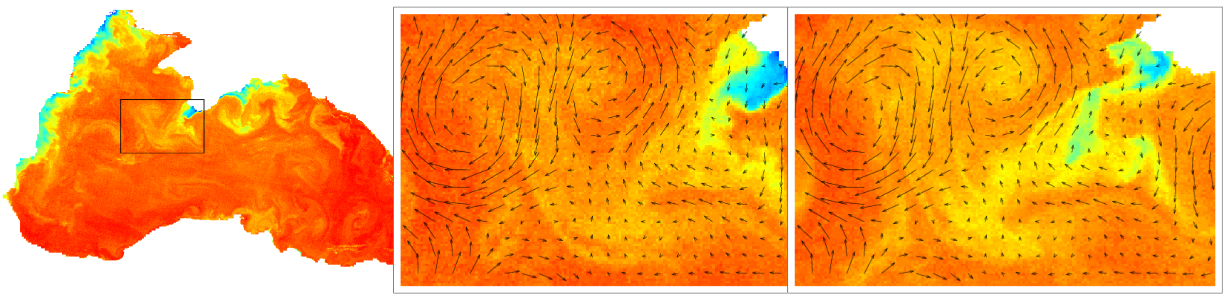


Figure 6. Zoom on a region of interest and motion estimation superposed on two consecutive images.

This research concerns the design of an operational method for rainfall nowcasting that aims at prevention of flash floods. The nowcasting method is based on two main components:

- a data assimilation method, based on radar images, estimates the state of the atmosphere: this is the estimation phase.
- a forecast method uses this estimation to extrapolate the state of the atmosphere in the future: this is the forecast phase.

Results were analyzed by Numtech (partner of a joint I-lab) on space-time neighborhood in order to prevent consequences of flash floods on previously defined zone.

Current research concerns the use of object components in the state vector in order to get an improved motion estimation and a better localization of endangered regions.

6.5. Uncertainty quantification and risk assessment

The uncertainty quantification of environmental models raises a number of problems due to:

- the dimension of the inputs, which can easily be 10^5-10^8 at every time step;
- the dimension of the state vector, which is usually 10^5-10^7;
- the high computational cost required when integrating the model in time.

While uncertainty quantification is a very active field in general, its implementation and development for geosciences requires specific approaches that are investigated by Clime. The project-team tries to determine the best strategies for the generation of ensembles of simulations. In particular, this requires addressing the generation of large multimodel ensembles and the issue of dimension reduction and cost reduction. The dimension reduction consists in projecting the inputs and the state vector to low-dimensional subspaces. The cost reduction is carried out by emulation, i.e., the replacement of costly components with fast surrogates.

6.5.1. Application of sequential aggregation to meteorology

Participants: Jean Thorey, Paul Baudin, Vivien Mallet, Stéphanie Dubost [EDF R&D], Christophe Chaussin [EDF R&D], Laurent Dubus [EDF R&D], Luc Musson-Genon [CEREA, EDF R&D], Laurent Descamps [Météo France], Philippe Blanc [Armines], Gilles Stoltz [CNRS].

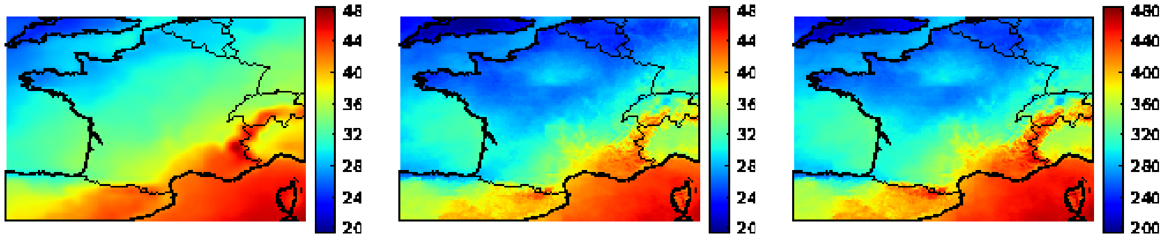
Nowadays, it is standard procedure to generate an ensemble of simulations for a meteorological forecast. Usually, meteorological centers produce a single forecast, out of the ensemble forecasts, computing the ensemble mean (where every model receives an equal weight). It is however possible to apply aggregation methods. When new observations are available, the meteorological centers also compute analyses. Therefore, we can apply the ensemble forecast of analyses. Ensembles of forecasts for mean sea level pressure, from the THORPEX Interactive Grand Global Ensemble, were aggregated with a forecast error decrease by 20% compared to the ensemble mean.

We studied the aggregation of ensembles of solar radiations in the context of photovoltaic production. The observations are based on MeteoSat Second Generation (MSG) and provided by the HelioClim-3 database as gridded fields. The ensembles of forecasts are from the THORPEX Interactive Grand Global Ensemble. The aggregated forecasts show a 20% error decrease compared to the individual forecasts. They are also able to retrieve finer spatial patterns than the ones found in the individual forecasts (see Figure 7).

6.5.2. Sequential aggregation with uncertainty estimation

Participants: Vivien Mallet, Jean Thorey, Paul Baudin, Gilles Stoltz [CNRS].

An important issue is the estimation of the uncertainties associated with the aggregated forecasts. We devised a new approach to predict a probability density function or cumulative distribution function instead of a single aggregated forecast. In practice, the aggregation procedure aims at forecasting the cumulative distribution function of the observations which is simply a Heaviside function centered at the observed value. Our forecast is the weighted empirical cumulative distribution function based on the ensemble of forecasts. The method guarantees that, in the long run, the forecast cumulative distribution function has a continuous ranked probability score at least as good as the best weighted empirical cumulative function with weights constant in time.



(a) (b) (c)

Figure 7. Yearly average of the map of downward shortwave solar radiation in $\mathrm{W m^{-2}}$*, for an ensemble mean (a), for our aggregated forecasts (b) and observed (c).*

6.5.3. Sensitivity analysis in the dispersion of radionuclides

Participants: Sylvain Girard, Vivien Mallet, Irène Korsakissok [IRSN].

We carried out a sensitivity analysis of the dispersion of radionuclides during Fukushima disaster. We considered the dispersion at regional scale, with the Eulerian transport model Polair3D from Polyphemus. The sensitivities to most input parameters were computed using the Morris method (with 8 levels and 100 trajectories). The influences of 19 scalar parameters were quantified. The scalar parameters were additive terms or multiplicative factors applied to 1D, 2D or 3D fields such as emission rates, precipitations, cloud height, wind velocity. The sensitivity analysis was carried out with the Morris method and by computing Sobol' indices. Both approaches were found to be consistent. Computing the Sobol' indices required the use of Gaussian process emulation, which proved to be successful at least on targets averaged in time and space.

It was shown that, depending on the output quantities of interest (various aggregated atmospheric and ground dose rates), the sensitivity to the inputs may greatly vary in time and space (see Figure 8). Very few parameters show low sensitivity in any case. The vertical diffusion coefficient, the scavenging factors, the winds and precipitation intensity were found to be the most influential inputs. Most input variables related to the source term (emission rates, emission dates) also had a strong influence.

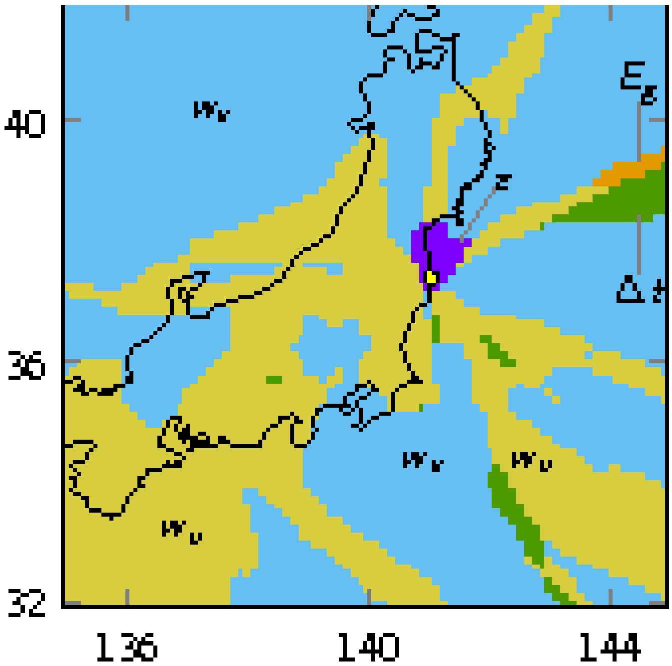


Figure 8. Variables that influence the most the atmospheric radioactivity after Fukushima disaster. z is the emissions altitude; Δt is the time shift on emissions; E_g stands for the emissions of noble gas; w_u and w_v are for zonal and meridional winds, respectively.

COAST Team

5. New Results

5.1. An authentication/authorization framework for federated environments

Participants: Ahmed Bouchami, Olivier Perrin.

Collaborative environments have put an enormous challenge on the security of information processing systems used to manage them. In the context of the Open PaaS project, we worked on a decentralised hybrid framework for managing access control designed for support of these environments. In our proposal, we manage thress dimensions: the authentication, the access control, and the governance of the security.

Our authentication framework supports an interoperable authentication, a combination of RBAC, XACML for decentralized multiple administration (authorization). Both identities and resources are federated: the former are controlled by PaaS Federated Security Modules, while the later are by a PaaS Federated Security Modules. This work has been presented in the I-ESA conference ([10]).

We have also proposed a formal cloud-based authorization framework. We have considered trust to be a dynamic attribute to facilitate authorization decisions and have proposed models to handle different qualitative, quantitative and periodicity based temporal constraints. Further, we have presented an architecture for policies evaluation in the cloud. We presented our model in the CollaborateCom conference [17]. The model relies on a formal event-calculus based approach. We have introduced an architecture that considers different levels at which authorization policies can be specified and decisions can be taken and combines user level policies with the enterprise policies, and it considers real-time and dynamic environment changes (context), supports timed delegation, and the computation and specification of attributes based on trust. An implementation has been integrated in the Open PaaS platform.

A third aspect deals with the governance of the security aspects (mainly authorization). In this part, we have proposed to audit the various accesses to the resources, and we have proposed a model which is able to lower/raise the trust level of a member of the federated community.

During this year, we have also implemented and integrated the framework in the Open PaaS prototype, and all the code is now accessible in the repository of the project. The integration is done, and the other components of the project are now using the authentication/authorization component.

5.2. Experimental user studies for collaborative editing

Participants: Mehdi Ahmed-Nacer, François Charoy, Claudia-Lavinia Ignat, Gérald Oster, Pascal Urso.

With several tools to support collaborative editing such as Google Drive and Etherpad, the practice of colaborative editing is increasingly common, e.g., group note taking during meetings and conferences, and brainstorming activities. While collaborative editing tools meet technical goals, the requirements for group performance are unclear. One system property of general interest is delay between a modification of a user is performed and this modification is visible to the other users. This delay can be caused by different reasons such as network delay due to physical communication technology, the complexity of various algorithms for ensuring consistency and the type of underlying architectures. No prior work questioned the maximum acceptable delay for real-time collaboration or the efficacy of compensatory strategies.

In [14] we studied the effect of delay on group performance on an artificial collaborative editing task where a group of four participants located the release dates for an alphabetized list of movies and re-sorted the list in chronological order. The experiment was performed with eighty users. We measured sorting accuracy based on the insertion sort algorithm, average time per entry, strategies (tightly coupled or loosely coupled task decomposition of the task) and chat behavior between users. We found out that delay slows down participants which decrements the outcome metric of sorting accuracy. Tightly coupled task decomposition enhances outcome at minimal delay, but participants slow down with higher delays. A loosely coupled task decomposition at the beginning leaves a poorly coordinated tightly coupled sorting at the end, requiring more coordination as delay increases.

In asynchronous collaborative editing, such as version control, the main feature to allow collaboration is the merge feature. However, software merging is a time-consuming and error-prone activity, and if a merge feature return results with too many conflicts and errors, this activity becomes even more difficult. To help developers, several algorithms have been proposed to improve the automation of merge tools. These algorithms aim at minimising conflict situations and therefore improving the productivity of the development team, however no general framework is proposed to evaluated and compare their result.

In [9] we propose a methodology to measure the effort required to use the result of a given merge tool. We employ the large number of publicly available open-source development histories to automatically compute this measure and evaluate the quality of the merging tools results. We use the simple idea that these histories contains both the concurrent modifications and their merge results as approved by the developers. Through a study of six open-source repositories totalling more than 2.5 millions lines of code, we show meaningful comparison results between merge algorithms and how to use the results to improve them.

5.3. Optimization and security of business processes in SaaS contexts

Participants: Claude Godart, Elio Goettelmann, Samir Youcef.

Globalization and the increase of competitive pressures created the need for agility in business processes, including the ability to outsource, offshore, to take opportunity of the cloud, or otherwise distribute its once-centralized business processes or parts thereof. While hampered thus far by limited infrastructure capabilities, the increase in bandwidth and connectivity and decrease in communication cost have removed these limits. This is even more true with the advent of cloud, particularly in its “Service as a software” dimension. To adapt to such a context, there is a growing need for the ability to fragment one’s business processes in an agile manner, and be able to distribute and wire these fragments so that their combined execution recreates the function of the original process. Our work is focused on solving some of the core challenges resulting from the need to dynamically restructure enterprise interactions. Restructuring such interactions corresponds to the fragmentation of intra- and inter-enterprise business process models. It describes how to identify, create, and execute process fragments without loosing the operational semantics of the original process models. In addition, this fragmentation is complicated by the constraints of quality of service, in particular the execution time and the cost, and of security, especially privacy. During the year, we consider this problem at two levels: the design of privacy-aware process models, and the optimization of process schedules. We developed a methodology to integrate privacy concerns in the design of a business process before distribution in the cloud [11]. Based on a risk analysis, the result of the design is a set of process (re)modeling actions, a set of constraints on process fragments assignments to clouds, and a set of constraints for cloud selection based on cloud properties [12].

COATI Project-Team

6. New Results

6.1. Network Design and Management

Participants: Jean-Claude Bermond, David Coudert, Frédéric Giroire, Frédéric Havet, Alvinice Kodjo, Aurélien Lancin, Bi Li, Fatima Zahra Moataz, Christelle Molle-Caillouet, Joanna Moulierac, Nicolas Nisse, Stéphane Pérennes, Truong Khoa Phan.

More information on several results presented in this section may be found in the PhD thesis of A. Kodjo [13], B. Li [15] and T. K. Phan [18].

6.1.1. Optimization in backbone networks

6.1.1.1. Shared Risk Link Group

The notion of Shared Risk Link Groups (SRLG) captures survivability issues when a set of links of a network may fail simultaneously. The theory of survivable network design relies on basic combinatorial objects that are rather easy to compute in the classical graph models: shortest paths, minimum cuts, or pairs of disjoint paths. In the SRLG context, the optimization criterion for these objects is no longer the number of edges they use, but the number of SRLGs involved. Unfortunately, computing these combinatorial objects is NP-hard and hard to approximate with this objective in general. Nevertheless some objects can be computed in polynomial time when the SRLGs satisfy certain structural properties of locality which correspond to practical ones, namely the star property (all links affected by a given SRLG are incident to a unique node) and the span 1 property (the links affected by a given SRLG form a connected component of the network). The star property is defined in a multi-colored model where a link can be affected by several SRLGs while the span property is defined only in a mono-colored model where a link can be affected by at most one SRLG. In [52], we extend these notions to characterize new cases in which these optimization problems can be solved in polynomial time or are fixed parameter tractable. We also investigate on the computational impact of the transformation from the multi-colored model to the mono-colored one. Experimental results are presented to validate the proposed algorithms and principles.

6.1.1.2. Dynamic Routing and Spectrum Assignment in Optical Networks

Elastic Optical Networks (EONs) promises a better utilization of the spectrum in optical networks. In fact, as the optical transmission spectrum is carved into fixed-length bands in the traditional WDM networks, small bit rates are over-provisioned and very high bit rates do not fit. EONs are moving away from this fixed-grid and allow the spectrum to be divided flexibly: each request is allocated exactly the resources it needs. In [34], we present two exact algorithms to route and allocate spectrum to a new request in an EON using only Non-Disruptive Defragmentation (Push-Pull). In the first algorithm, we find the shortest routing path for the new request (i.e., the shortest path from source to destination where contiguous spectrum to satisfy the request can be freed) and then find the position that gives the overall minimum delay on that path. In the second algorithm, we find at the same time a routing path and a position in the spectrum, that minimize the delay of insertion (over all other paths and positions). Both algorithms are polynomial in the size of the network, its bandwidth and the number of provisioned requests.

6.1.2. Microwave Backhaul networks

6.1.2.1. Chance-Constrained Optimization of Reliable Backhaul networks

In [25], we extend our former investigation on conceiving reliable fixed point-to-point wireless networks under outage probability constraints. We consider the problem of determining the minimum cost bandwidth assignment of a network, while guaranteeing a reliability level of the solution. If the optimal bandwidth assignment and routing of traffic demands are accomplished, the reliability criterion requires that network flows remain feasible with high probability, regarding that the performance of microwave links is prone to variations due to external factors, e.g., weather. We introduce a *chance-constrained programming* approach to tackle this problem and we present reformulations to standard integer linear programming models, including a budget-constrained formulation. To improve the solving performance, we propose new valid inequalities and a primal heuristic. Computational results present a performance analysis of the valid inequalities and the heuristic. Further, the outperformance of the novel model compared to more traditional approaches is documented.

6.1.2.2. Robust optimization in multi-operators microwave backhaul networks

In [41], we consider the problem of sharing the infrastructure of a backhaul network for routing. We investigate on the revenue maximization problem for the physical network operator (PNO) when subject to stochastic traffic requirements of multiple virtual network operators (VNO) and prescribed service level agreements (SLA). We use robust optimization to study the tradeoff between revenue maximization and the allowed level of uncertainty in the traffic demands. This mixed integer linear programming model takes into account end-to-end traffic delays as example of quality-of-service requirement in a SLA. To show the effectiveness of our model, we present a study on the price of robustness, i.e. the additional price to pay in order to obtain a feasible solution for the robust scheme, on realistic scenarios.

6.1.3. Energy efficiency

6.1.3.1. Robust Optimization for Energy-aware Routing with Redundancy Elimination

Many studies in literature have shown that energy-aware routing (EAR) can significantly reduce energy consumption for backbone networks. Also, as an arising concern in networking research area, the protocol-independent traffic redundancy elimination (RE) technique helps to reduce (a.k.a compress) traffic load on backbone network. In [35], [50], we present an extended model of the classical multi-commodity flow problem with compressible flows. Our model is robust with fluctuation of traffic demand and compression rate. In details, we allow any set of a predefined size of traffic flows to deviate simultaneously from their nominal volumes or compression rates. As an applicable example, we use this model to combine redundancy elimination and energy-aware routing to increase energy efficiency for a backbone network. Using this extra knowledge on the dynamics of the traffic pattern, we are able to significantly increase energy efficiency for the network. We formally define the problem and model it as a Mixed Integer Linear Program (MILP). We then propose an efficient heuristic algorithm that is suitable for large networks. Simulation results with real traffic traces on Abilene, Geant and Germany50 networks show that our approach allows for 16-28% extra energy savings with respect to the classical EAR model.

6.1.3.2. Optimizing IGP Link Weights for Energy-efficiency in a Changing World

Recently, due to the increasing power consumption and worldwide gas emissions in ICT (Information and Communication Technology), energy efficient ways to design and operate backbone networks are becoming a new concern for network operators. Since these networks are usually overprovisioned and since traffic load has a small influence on power consumption of network equipments, the most common approach to save energy is to put unused line cards that drive links between neighbouring routers into sleep mode. To guarantee QoS, all traffic demands should be routed without violating capacity constraints and the network should keep its connectivity. From the perspective of traffic engineering, we argue that stability in routing configuration also plays an important role in QoS. In details, frequent changes in network configuration (link weights, slept and activated links) to adapt with traffic fluctuation in daily time cause network oscillations. We propose in [62] a novel optimization method to adjust the link weights of Open Shortest Path First (OSPF) protocol while limiting the changes in network configurations when multi-period traffic matrices are considered. We formally define the problem and model it as Mixed Integer Linear Program (MILP). We then propose an efficient heuristic algorithm that is suitable for large networks. Simulation results with real traffic traces on three different networks show that our approach achieves high energy saving while keeping the networks in stable state (less changes in network configuration).

6.1.3.3. Grid spanners with low forwarding index for energy efficient networks

A routing R of a connected graph G is a collection that contains simple paths connecting every ordered pair of vertices in G. The edge-forwarding index with respect to R (or simply the forwarding index with respect to R) $\pi(G, R)$ of G is the maximum number of paths in R passing through any edge of G. The forwarding index $\pi(G)$ of G is the minimum $\pi(G, R)$ over all routings R's of G. This parameter has been studied for different graph classes. Motivated by energy efficiency, we look in [57], for different numbers of edges, at the best spanning graphs of a square grid, namely those with a low forwarding index.

6.1.4. Software-Defined Networks

6.1.4.1. Rule Placement in Software-Defined Networks for Energy-aware Routing

Software-defined Networks (SDN), in particular OpenFlow, is a new networking paradigm enabling innovation through network programmability. Over past few years, many applications have been built using SDN such as server load balancing, virtual-machine migration, traffic engineering and access control. We focus on using SDN for energy-aware routing (EAR). SDN can collect traffic matrix and then computes routing solutions satisfying QoS while being minimal in energy consumption (with minimal number of active links). However, prior works on EAR have assumed that the table of OpenFlow switch can hold an infinite number of rules. In practice, this assumption does not hold since the flow table is implemented with Ternary Content Addressable Memory (TCAM) which is expensive and power-hungry. In [39], [56], we propose an optimization method to minimize energy consumption for a backbone network while respecting capacity constraints on links and rule space constraints on routers. In details, we present an exact formulation using Integer Linear Program (ILP) and introduce efficient greedy heuristic algorithm. Based on simulations, we show that using this smart rule space allocation, it is possible to save almost as much power consumption as the classical EAR approach.

6.1.4.2. Compressing Two-dimensional Routing Tables with Order

A communication in a network is a pair of nodes (s, t). The node s is called the source source and t the destination. A communication set is a set of distinct communications, i.e. two communications might have the same source or the same destination, but they cannot have both same source and same destination. A routing of a communication (s, t) is a path in the network from s to t. A routing of a communication set is a union of routings of its communications. At each node, there is a set X of communications whose routing path goes through this node. The node needs to be able to find for each communication (s, t) in X, the port that the routing path of (s, t) uses to leave it. An easy way of doing it is to store the list of all triples (s, t, k), where $(s, t) \in X$ and k is the port used by the (s, t)-path to leave the node. Such triples are called communication triples. However, such a list might be very large. Motivated by routing in telecommunication network using Software Defined Network Technologies, we consider in [55] the problem of compacting this list using aggregation rules. Indeed, SDN routers use specific memory which is expensive and of small capacity. Hence, in addition, we can use some additional triples, called *-triples. As an example, a t-destination triple $(*, t, p)$, means that every communication with destination t leaves on port p. We carry out in this work a study of the problem complexity, providing results of NP-completeness, of Fixed-Parameter Tractability and approximation algorithms.

6.1.5. Data gathering in radio networks

In the gathering problem, a particular node in a graph, the base station, aims at receiving messages from some nodes in the graph. At each step, a node can send one message to one of its neighbors (such an action is called a call). However, a node cannot send and receive a message during the same step. Moreover, the communication is subject to interference constraints ; more precisely we consider a binary interference model where two calls interfere in a step, if the sender of one call is at distance at most d_I from the receiver of the other call. Given a graph with a base station and a set of nodes having some messages, the goal of the gathering problem is to compute a schedule of calls for the base station to receive all messages as fast as possible, i.e., minimizing the number of steps (called makespan). The gathering problem is equivalent to the personalized broadcasting problem where the base station has to send messages to some nodes in the graph, with same transmission constraints. In [23], we focus on the gathering and personalized broadcasting problem in grids. Moreover, we consider the non-buffering model: when a node receives a message at some step, it must transmit it during the next step. In this setting, though the problem of determining the complexity of computing the optimal makespan in a grid is still open, we present linear (in the number of messages) algorithms that compute schedules for gathering with $d_I \in \{0,1,2\}$. In particular, we present an algorithm that achieves the optimal makespan up to an additive constant 2 when $d_I = 0$. If no messages are "close" to the axes (the base station being the origin), our algorithms achieve the optimal makespan up to an additive constant 1 when $d_I = 0$, 4 when $d_I = 2$, and 3 when both $d_I = 1$ and the base station is in a corner. Note that, the approximation algorithms that we present also provide approximation up to a ratio 2 for the gathering with buffering. All our results are proved in terms of personalized broadcasting.

6.2. Graph Algorithms

Participants: Julio Araújo, Jean-Claude Bermond, David Coudert, Guillaume Ducoffe, Frédéric Giroire, Aurélien Lancin, Bi Li, Fatima Zahra Moataz, Christelle Molle-Caillouet, Nicolas Nisse, Stéphane Pérennes.

COATI is also interested in the algorithmic aspects of Graph Theory. In general we try to find the most efficient algorithms to solve various problems of Graph Theory and telecommunication networks. More information on several results presented in this section may be found in PhD thesis of B. Li [15] and A. Lancin [14], and in the Habilitation thesis of N. Nisse [17].

6.2.1. Complexity and Computation of Graph Parameters

We use graph theory to model various network problems. In general we study their complexity and then we investigate the structural properties of graphs that make these problems hard or easy. In particular, we try to find the most efficient algorithms to solve the problems, sometimes focusing on specific graph classes from which the problems are polynomial-time solvable.

6.2.1.1. Hyperbolicity

The Gromov hyperbolicity is an important parameter for analyzing complex networks since it expresses how the metric structure of a network looks like a tree. In other words, it provides bounds on the stretch resulting from the embedding of a network topology into a weighted tree. It is therefore used to provide bounds on the expected stretch of greedy-routing algorithms in Internet-like graphs. However, the best known algorithm for computing this parameter has time complexity in $O(n^{3.69})$, which is prohibitive for large-scale graphs.

In [47], we investigate some relations between the hyperbolicity of a graph and the hyperbolicity of its *atoms*, that are the subgraphs resulting from the decomposition of the graph according its clique minimal separators. More precisely, we prove that the maximum hyperbolicity taken over all the atoms is at least the hyperbolicity of *Gminus one*. We also give an algorithm to slightly modify the atoms, which is at no extra cost than computing the atoms themselves, and so that the maximum hyperbolicity taken over all the resulting graphs is *exactly* the hyperbolicity of G. An experimental evaluation of our methodology is provided for large collaboration networks. Finally, we deduce from our theoretical results the first *linear-time* algorithm to compute the hyperbolicity of an outerplanar graph.

The shortest-path metric d of a connected graph G is 1/2-hyperbolic if, and only if, it satisfies $\mathrm{d}(u,v)+\mathrm{d}(x,y) \leq \max\{\mathrm{d}(u,x)+\mathrm{d}(v,y), \mathrm{d}(u,y)+\mathrm{d}(v,x)\}+1$, for every 4-tuple u,x,v,y of G. We show in [26], [48] that the problem of deciding whether an unweighted graph is 1/2-hyperbolic is subcubic equivalent to the problem of determining whether there is a chordless cycle of length 4 in a graph. An improved algorithm is also given for both problems, taking advantage of fast rectangular matrix multiplication. In the worst case it runs in $O(n^{3.26})$-time.

6.2.1.2. Branch and Bound Algorithm for computing Pathwidth

It is well known that many NP-hard problems are tractable in the class of bounded pathwidth graphs. In particular, path-decompositions of graphs are an important ingredient of dynamic programming algorithms for solving such problems. Therefore, computing the pathwidth and associated path-decomposition of graphs has both a theoretical and practical interest. In [36], [51], we design a Branch and Bound algorithm that computes the exact pathwidth of graphs and a corresponding path-decomposition. Our main contribution consists of several non-trivial techniques to reduce the size of the input graph (pre-processing) and to cut the exploration space during the search phase of the algorithm. We evaluate experimentally our algorithm by comparing it to existing algorithms of the literature. It appears from the simulations that our algorithm offers a significative gain with respect to previous work. In particular, it is able to compute the exact pathwidth of any graph with less than 60 nodes in a reasonable running-time (10 min.). Moreover, our algorithm also achieves good performance when used as a heuristic (i.e., when returning best result found within bounded time-limit). Our algorithm is not restricted to undirected graphs since it actually computes the vertex-separation of digraphs (which coincides with the pathwidth in case of undirected graphs).

6.2.1.3. To satisfy impatient Web surfers is hard

Prefetching is a basic mechanism for faster data access and efficient computing. An important issue in prefetching is the tradeoff between the amount of network's resources wasted by the prefetching and the gain of time. For instance, in the Web, browsers may download documents in advance while a Web surfer is surfing. Since the Web surfer follows the hyperlinks in an unpredictable way, the choice of the Web pages to be prefetched must be computed online. The question is then to determine the minimum amount of resources used by prefetching that ensures that all documents accessed by the Web surfer have previously been loaded in the cache. In [28], we model this problem as a two-player game similar to Cops and Robber Games in graphs. Let $k \geq 1$ be any integer. The first player, a fugitive, starts on a marked vertex of a (di)graph G. The second player, an observer, marks at most k vertices, then the fugitive moves along one edge/arc of G to a new vertex, then the observer marks at most k vertices, etc. The fugitive wins if it enters an unmarked vertex, and the observer wins otherwise. The surveillance number of a (di)graph is the minimum k such that the observer marking at most k vertices at each step can win against any strategy of the fugitive. We also consider the connected variant of this game, i.e., when a vertex can be marked only if it is adjacent to an already marked vertex. We study the computational complexity of the game. All our results hold for both variants, connected or unrestricted. We show that deciding whether the surveillance number of a chordal graph is at most 2 is NP-hard. We also prove that deciding if the surveillance number of a DAG is at most 4 is PSPACE-complete. Moreover, we show that the problem of computing the surveillance number is NP-hard in split graphs. On the other hand, we provide polynomial-time algorithms computing surveillance numbers of trees and interval graphs. Moreover, in the case of trees, we establish a combinatorial characterization of the surveillance number.

6.2.2. Tree-decompositions

6.2.2.1. Minimum Size Tree-Decompositions

Tree-Decompositions are the corner-stone of many dynamic programming algorithms for solving graph problems. Since the complexity of such algorithms generally depends exponentially on the width (size of the bags) of the decomposition, much work has been devoted to compute tree-decompositions with small width. However, practical algorithms computing tree-decompositions only exist for graphs with treewidth less than 4. In such graphs, the time-complexity of dynamic programming algorithms based on tree-decompositions is dominated by the size (number of bags) of the tree-decompositions. It is then interesting to try to minimize the size of the tree-decompositions. In [42], [60], we consider the problem of computing a tree-decomposition of a graph with width at most k and minimum size. More precisely, we focus on the following problem: given a fixed $k \geq 1$, what is the complexity of computing a tree-decomposition of width at most k with minimum size in the class of graphs with treewidth at most k? We prove that the problem is NP-complete in planar graphs for any fixed $k \geq 4$ and polynomial for $k \leq 2$. We also show that for $k = 3$ the problem can be solved in polynomial time in the class of trees and 2-connected outerplanar graphs.

6.2.2.2. Exclusive Graph Searching vs. Pathwidth

In Graph Searching, a team of searchers aims at capturing an invisible fugitive moving arbitrarily fast in a graph. Equivalently, the searchers try to clear a contaminated network. The problem is to compute the minimum number of searchers required to accomplish this task. Several variants of Graph Searching have been studied mainly because of their close relationship with the pathwidth of a graph. Blin et al. defined the Exclusive Graph Searching where searchers cannot "jump" and no node can be occupied by more than one searcher. In [61], we study the complexity of this new variant. We show that the problem is NP-hard in planar graphs with maximum degree 3 and it can be solved in linear time in the class of cographs. We also show that monotone Exclusive Graph Searching is NP-complete in split graphs where Pathwidth is known to be solvable in polynomial time. Moreover, we prove that monotone Exclusive Graph Searching is in P in a subclass of star-like graphs where Pathwidth is known to be NP-hard. Hence, the computational complexities of monotone Exclusive Graph Searching and Pathwidth cannot be compared. This is the first variant of Graph Searching for which such a difference is proved.

6.2.2.3. Diameter of Minimal Separators in Graphs

In [49], we establish general relationships between the topological properties of graphs and their metric properties. For this purpose, we upper-bound the diameter of the *minimal separators* in any graph by a function of their sizes. More precisely, we prove that, in any graph G, the diameter of any minimal separator S in G is at most $\lfloor\frac{\ell(G)}{2}\rfloor \cdot (|S|-1)$ where $\ell(G)$ is the maximum length of an isometric cycle in G. We refine this bound in the case of graphs admitting a *distance preserving ordering* for which we prove that any minimal separator S has diameter at most $2(|S|-1)$. Our proofs are mainly based on the property that the minimal separators in a graph G are connected in some power of G.

Our result easily implies that the *treelengthtl*(G) of any graph G is at most $\lfloor\frac{\ell(G)}{2}\rfloor$ times its *treewidthtw*(G). In addition, we prove that, for any graph G that excludes an *apex graphH* as a minor, $tw(G) \leq c_H \cdot tl(G)$ for some constant c_H only depending on H. We refine this constant when G has bounded genus. As a consequence, we obtain a very simple $O(\ell(G))$-approximation algorithm for computing the treewidth of n-node m-edge graphs that exclude an apex graph as a minor in $O(nm)$-time.

6.2.3. Distributed computing with mobile agents

6.2.3.1. Stigmergy of Anonymous Agents in Discrete Environments

Communication by stigmergy consists, for agents/robots devoid of other dedicated communication devices, in exchanging information by observing each other's movements, similar to how honeybees use a dance to inform each other on the location of food sources. Stigmergy, while a popular technique in soft computing (e.g., swarm intelligence and swarm robotics), has received little attention from a computational viewpoint, with only one study proposing a method in a continuous environment. An important question is whether there are limits intrinsic to the environment on the feasibility of stigmergy. While it is not the case in a continuous environment, we show that the answer is quite different when the environment is discrete. In [53], [37], we consider stigmergy in graphs and identifies classes of graphs in which robots can communicate by stigmergy. We provide two algorithms with different tradeoffs. One algorithm achieves faster stigmergy when the density of robots is low enough to let robots move independently. This algorithm works when the graph contains some particular pairwise-disjoint subgraphs. The second algorithm, while slower solves the problem under an extremely high density of robots assuming that the graph admits some large cycle. Both algorithms are described in a general way, for any graph that admits the desired properties and with identified nodes. We show how the latter assumption can be removed in more specific topologies. Indeed, we consider stigmergy in the grid which offers additional orientation information not available in a general graphs, allowing us to relax some of the assumptions. Given an $N \times M$ anonymous grid, we show that the first algorithm requires $O(\mathcal{M})$ steps to achieve communication by stigmergy, where $\mathcal{M}$ is the maximum length of a communication message, but it works only if the number of robots is less than $\lfloor\frac{N \cdot M}{9}\rfloor$. The second algorithm, which requires $O(k^2)$ steps, where k is the number of robots, on the other hand, works for up to $N \cdot M - 5$ robots. In both cases, we consider very weak assumptions on the robots capabilities: i.e., we assume that the robots are anonymous, asynchronous, uniform, and execute deterministic algorithms.

6.2.3.2. Gathering and Exclusive Searching on Rings under Minimal Assumptions

Consider a set of mobile robots with minimal capabilities placed over distinct nodes of a discrete anonymous ring. Asynchronously, each robot takes a snapshot of the ring, determining which nodes are either occupied by robots or empty. Based on the observed configuration, it decides whether to move to one of its adjacent nodes or not. In the first case, it performs the computed move, eventually. The computation also depends on the required task. In [38], we solve both the well-known Gathering and Exclusive Searching tasks. In the former problem, all robots must simultaneously occupy the same node, eventually. In the latter problem, the aim is to clear all edges of the graph. An edge is cleared if it is traversed by a robot or if both its endpoints are occupied. We consider the exclusive searching where it must be ensured that two robots never occupy the same node. Moreover, since the robots are oblivious, the clearing is perpetual, i.e., the ring is cleared infinitely often. In the literature, most contributions are restricted to a subset of initial configurations. Here, we design two different algorithms and provide a characterization of the initial configurations that permit the resolution of the problems under minimal assumptions.

6.2.4. Enhancing the Web's Transparency

Today's Web services – such as Google, Amazon, and Facebook – leverage user data for varied purposes, including personalizing recommendations, targeting advertisements, and adjusting prices. At present, users have little insight into how their data is being used. Hence, they cannot make informed choices about the services they choose.

To increase transparency, we developed *XRay* [40], the first fine-grained, robust, and scalable personal data tracking system for the Web. XRay predicts which data in an arbitrary Web account (such as emails, searches, or viewed products) is being used to target which outputs (such as ads, recommended products, or prices). XRay's core functions are service agnostic and easy to instantiate for new services, and they can track data within and across services. To make predictions independent of the audited service, XRay relies on the following insight: by comparing outputs from different accounts with similar, but not identical, subsets of data, one can pinpoint targeting through correlation. We show both theoretically, and through experiments on Gmail, Amazon, and YouTube, that XRay achieves high precision and recall by correlating data from a surprisingly small number of extra accounts.

6.2.5. Algorithm design in biology

In COATI, we have recently started a collaboration with EPI ABS (Algorithms Biology Structure) from Sophia Antipolis on minimal connectivity complexes in mass spectrometry based macro-molecular complex reconstruction [63]. This problem turns out to be a minimum color covering problem (minimum number of colors to cover colored edges with connectivity constraints on the subgraphs induced by the colors) of the edges of a graph, and is surprizingly similar to a capacity maximization problem in a multi-interfaces radio network we were studying.

Consider a set of oligomers listing the subunits involved in sub-complexes of a macro-molecular assembly, obtained e.g. using native mass spectrometry or affinity purification. Given these oligomers, connectivity inference (CI) consists of finding the most plausible contacts between these subunits, and minimum connectivity inference (MCI) is the variant consisting of finding a set of contacts of smallest cardinality. MCI problems avoid speculating on the total number of contacts, but yield a subset of all contacts and do not allow exploiting a priori information on the likelihood of individual contacts. In this context, we present in [43] two novel algorithms, ALGO-MILP-W and ALGO-MILP-WB. The former solves the minimum weight connectivity inference (MWCI), an optimization problem whose criterion mixes the number of contacts and their likelihood. The latter uses the former in a bootstrap fashion, to improve the sensitivity and the specificity of solution sets. Experiments on the yeast exosome, for which both a high resolution crystal structure and a large set of oligomers is known, show that our algorithms predict contacts with high specificity and sensitivity, yielding a very significant improvement over previous work. The software accompanying this paper is made available, and should prove of ubiquitous interest whenever connectivity inference from oligomers is faced.

6.3. Structural Graph Theory

Participants: Jean-Claude Bermond, Frédéric Havet, Nicolas Nisse, Ana Karolinna Maia de Oliveira, Stéphane Pérennes.

More information on several results presented in this section may be found in PhD thesis of A. K. Maia de Oliveira [16], and in the Habilitation thesis of N. Nisse [17].

6.3.1. Graph colouring and applications

Graph colouring is a central problem in graph theory and it has a huge number of applications in various scientific domains (telecommunications, scheduling, bio-informatics, ...). We mainly study graph colouring problems that model ressource allocation problems.

6.3.1.1. Backbone colouring

A well-known channel assignment problem is the following: we are given a graph G, whose vertices correspond to transmitters, together with an edge-weighting w. The weight of an edge corresponds to the minimum separation between the channels on its endvertices to avoid interferences. (If there is no edge, no separation is required, the transmitters do not interfere.) We need to assign positive integers (corresponding to channels) to the vertices so that for every edge e the channels assigned to its endvertices differ by at least $w(e)$. The goal is to minimize the largest integer used, which corresponds to minimizing the *span* of the used bandwidth. We studied a particular, yet quite general, case, called *backbone colouring*, in which there are only two levels of interference. So we are given a graph G and a subgraph H, called *the backone*. Two adjacent vertices in H must get integers at least q apart, while adjacent vertices in G must get integers at distance at least 1. The minimum span in this case is called the q-backbone chromatic number and is denoted $BBC_q(G,H)$. In [30] and [45], we focus on the case when G is planar and H is a forest. In [30], we give a series of NP-hardness results as well as upper bounds for $BBC_q(G,H)$, depending on the type of the forest (matching, galaxy, spanning tree). We also discuss a circular version of the problem. In [45], we give some upper bounds when G is planar and has no cycles of length 4 an 5, and G is a tree, and we relate those results to the celebrated Steinberg's Conjecture stating that every planar graphs with no cycles of length 4 or 5 is 3-colourable.

In [29], we consider the list version of this problem (in which each vertex is given a particular list of admissible colours), with particular focus on colours in $\mathbb{Z}_p$ – this problem is closely related to the problem of circular choosability. We first prove that the list circular q-backbone chromatic number of a graph is bounded by a function of the list chromatic number. We then consider the more general problem in which each edge is assigned an individual distance between its endpoints, and provide bounds using the Combinatorial Nullstellensatz. Through this result and through structural approaches, we achieve good bounds when both the graph and the backbone belong to restricted families of graphs.

6.3.1.2. On-line colouring graphs with few P_4s

Various on-line colouring procedures are used. The most widespread ones is the greedy one, which results in a greedy colouring. Given a graph $G = (V;E)$, a *greedy colouring* of G is a proper colouring such that, for each two colours $i < j$, every vertex of $V(G)$ coloured j has a neighbour with colour i. A second optimization procedure consists from time to time to consider the present colouring and to free some colour when possible: if each vertex of a colour class has another colour that is not used by its neighbours, we can recolour each vertex in the calls by another colour. This procedure results in a b-colouring of the graph. A *b-colouring* of a graph G is a proper colouring such that every colour class contains a vertex which is adjacent to at least one vertex in every other colour class. One of the performance measure of such graph is the maximum numbers of colours they could possibly use. The greatest k such that G has a greedy colouring with k colours is the *Grundy number* of G. The greatest integer k for which there exists a b-colouring of G with k colours is its *b-chromatic number*. Determining the Grundy number and the b-chromatic number of a graph are NP-hard problems in general. For a fixed q, the $(q;q-4)$-graphs are the graphs for which no set of at most q vertices induces more than $q-4$ distinct induced P_4s paths of order 4). In [24], we obtain polynomial-time algorithms to determine the Grundy number and the b-chromatic number of $(q;q-4)$-graphs, for a fixed q. They generalize previous results obtained for cographs and P_4-sparse graphs, classes strictly contained in the $(q;q-4)$-graphs.

6.3.1.3. Weighted colouring

We also studied weighted colouring which models various problems of shared resources allocation. Given a vertex-weighted graph G and a (proper) r-colouring $c = \{C_1, ..., C_r\}$ of G , the weight of a colour class C_i is the maximum weight of a vertex coloured i and the weight of c is the sum of the weights of its colour classes. The objective of the Weighted Colouring Problem is, given a vertex-weighted graph G , to determine the minimum weight of a proper colouring of G, that is, its *weighted chromatic number*. In [21], [33], we prove that the Weighted Coloring Problem admits a version of the Hajós' Theorem and so we show a necessary and sufficient condition for the weighted chromatic number of a vertex-weighted graph G to be at least k , for any positive real k. The Weighted Colouring Problem problem remains NP-complete in some particular graph classes as bipartite graphs. In their seminal paper, Guan and Zhu asked whether the weighted chromatic number of bounded tree-width graphs (partial k-trees) can be computed in polynomial-time. Surprisingly, the time-complexity of computing this parameter in trees is still open. We show in [21] that, assuming the Exponential Time Hypothesis (3-SAT cannot be solved in sub-exponential time), the best algorithm to compute the weighted chromatic number of n-node trees has time-complexity $n^{\Theta(\log n)}$. Our result mainly relies on proving that, when computing an optimal proper weighted colouring of a graph G, it is hard to combine colourings of its connected components, even when G is a forest.

6.3.1.4. Inducing proper colourings

Frequently, the proper colouring of the graph must be induced by some other parameters that a vertex can compute locally, for example on looking on the labels assigned to its incident edges or to their orientations.

For a connected graph G of order $|V(G)| \geq 3$ and a k-labelling $c : E(G) \rightarrow \{1, 2, ..., k\}$ of the edges of G, the *code* of a vertex v of G is the ordered k-tuple $(\ell_1, \ell_2, ..., \ell_k)$, where ℓ_i is the number of edges incident with v that are labelled i. The k-labelling c is *detectable* if every two adjacent vertices of G have distinct codes. The minimum positive integer k for which G has a detectable k-labelling is the *detection number*$det(G)$ of G. In [31], we show that it is NP-complete to decide if the detection number of a cubic graph is 2. We also show that the detection number of every bipartite graph of minimum degree at least 3 is at most 2. Finally, we give some sufficient condition for a cubic graph to have detection number 3.

An *orientation* of a graph G is a digraph D obtained from G by replacing each edge by exactly one of the two possible arcs with the same endvertices. For each $v \in V(G)$, the *indegree* of v in D, denoted by $d^-_D(v)$, is the number of arcs with head v in D. An orientation D of G is *proper* if $d^-_D(u) \neq d^-_D(v)$, for all $uv \in E(G)$. The *proper orientation number* of a graph G, denoted by $po(G)$, is the minimum of the maximum indegree over all its proper orientations. In [32], [44], we prove that $po(G) \leq \left(\Delta(G) + \sqrt{\Delta(G)}\right)/2 + 1$ if G is a bipartite graph, and $po(G) \leq 4$ if G is a tree. It is well-known that $po(G) \leq \Delta(G)$, for every graph G. However, we prove that deciding whether $po(G) \leq \Delta(G) - 1$ is already an NP-complete problem on graphs with $\Delta(G) = k$, for every $k \geq 3$. We also show that it is NP-complete to decide whether $po(G) \leq 2$, for planar *subcubic* graphs G. Moreover, we prove that it is NP-complete to decide whether $po(G) \leq 3$, for planar bipartite graphs G with maximum degree 5.

6.3.2. Directed graphs

Graph theory can be roughly partitioned into two branches: the areas of undirected graphs and directed graphs (digraphs). Even though both areas have numerous important applications, for various reasons, undirected graphs have been studied much more extensively than directed graphs. One of the reasons is that many problems for digraphs are much more difficult than their analogues for undirected graphs.

6.3.2.1. Finding a subdivision of a digraph

One of the cornerstones of modern (undirected) graph theory is minor theory of Robertson and Seymour. Unfortunately, we cannot expect an equivalent for directed graphs. Minor theory implies in particular that, for any fixed F, detecting a subdivision of F in an input graph G can be performed in polynomial time by the Robertson and Seymour linkage algorithm. In contrast, the analogous subdivision problem for digraph can be either polynomial-time solvable or NP-complete, depending on the fixed digraph F. In [16], a number of examples of polynomial instances, several NP-completeness proofs as well as a number of conjectures and open problems are given. In addition, it is conjectured that, for every integer k greater than 1, the directed cycles of length at least k have the Erdős-Pósa Property : for every n, there exists an integer t_n such that for every digraph D, either D contains n disjoint directed cycles of length at least k, or there is a set T of t_n vertices that meets every directed cycle of length at least k. This generalizes a celebrated result of Reed, Robertson, Seymour and Thomas which is the case $k = 2$ of this conjecture. We prove the conjecture for $k = 3$. We also show that the directed k-Linkage problem is polynomial-time solvable for digraphs with circumference at most 2. From these two results, we deduce that if F is the disjoint union of directed cycles of length at most 3, then one can decide in polynomial time if a digraph contains a subdivision of F.

6.3.2.2. The complexity of finding arc-disjoint branching flows

The concept of arc-disjoint flows in networks is a very general framework within which many well-known and important problems can be formulated. In particular, the existence of arc-disjoint branching flows, that is, flows which send one unit of flow from a given source s to all other vertices, generalizes the concept of arc-disjoint out-branchings (spanning out-trees) in a digraph. A pair of out-branchings $B^+_{s,1}$, $B^+_{s,2}$ from a root s in a digraph $D = (V, A)$ on n vertices corresponds to arc-disjoint branching flows x_1,x_2 (the arcs carrying flow in x_i are those used in $B^+_{s,i}$, $i = 1, 2$) in the network that we obtain from D by giving all arcs capacity $n-1$. It is then a natural question to ask how much we can lower the capacities on the arcs and still have, say, two arc-disjoint branching flows from the given root s. In [46], we prove that for every fixed integer ≥ 2 it is

- an NP-complete problem to decide whether a network $\mathcal{N} = (V, A, u)$ where $u_{ij} = k$ for every arc ij has two arc-disjoint branching flows rooted at s.
- a polynomial problem to decide whether a network $\mathcal{N} = (V, A, u)$ on n vertices and $u_{ij} = n - k$ for every arc ij has two arc-disjoint branching flows rooted at s.

The algorithm for the later result generalizes the polynomial algorithm, due to Lovász, for deciding whether a given input digraph has two arc-disjoint out-branchings rooted at a given vertex. Finally we prove that under the so-called Exponential Time Hypothesis (ETH), for every $\epsilon > 0$ and for every $k(n)$ with $(\log (n))^{1+\epsilon} \leq k(n) \leq \frac{n}{2}$ (and for every large i we have $k(n) = i$ for some n) there is no polynomial algorithm for deciding whether a given digraph contains two arc-disjoint branching flows from the same root so that no arc carries flow larger than $n - k(n)$.

6.3.2.3. Splitting a tournament into two subtournaments with given minimum outdegree

A (k_1, k_2)-*outdegree-splitting* of a digraph D is a partition (V_1, V_2) of its vertex set such that $D[V_1]$ and $D[V_2]$ have minimum outdegree at least k_1 and k_2, respectively. In [58], we show that there exists a minimum function f_T such that every tournament of minimum outdegree at least $f_T(k_1, k_2)$ has a (k_1, k_2)-outdegree-splitting, and $f_T(k_1, k_2) \leq k_1^2/2 + 3k_1/2 + k_2 + 1$. We also show a polynomial-time algorithm that finds a (k_1, k_2)-outdegree-splitting of a tournament if one exists, and returns 'no' otherwise. We give better bound on f_T and faster algorithms when $k_1 = 1$.

6.3.2.4. Eulerian and Hamiltonian dicycles in directed hypergraphs

In [19], we generalize the concepts of Eulerian and Hamiltonian digraphs to directed hypergraphs. A *dihypergraph*H is a pair $(\mathcal{V}(H), \mathcal{E}(H))$, where $\mathcal{V}(H)$ is a non-empty set of elements, called *vertices*, and $\mathcal{E}(H)$ is a collection of ordered pairs of subsets of $\mathcal{V}(H)$, called *hyperarcs*. It is Eulerian (resp. Hamiltonian) if there is a dicycle containing each hyperarc (resp. each vertex) exactly once. We first present some properties of Eulerian and Hamiltonian dihypergraphs. For example, we show that deciding whether a dihypergraph is Eulerian is an NP-complete problem. We also study when iterated line dihypergraphs are Eulerian and Hamiltonian. Finally, we study when the generalized de Bruijn dihypergraphs are Eulerian and Hamiltonian. In particular, we determine when they contain a complete Berge dicycle, i.e. an Eulerian and Hamiltonian dicycle.

COFFEE Project-Team (section vide)

COMETE Project-Team

6. New Results

6.1. Highlights of the Year

- Prix de thèse de l'Ecole Polytechnique 2014 for the thesis "The Epistemic View of Concurrency Theory" by Sophia Knight (Defended 20 September, 2013).
- Catuscia Palamidessi has been invited keynote speaker at the joint conferences CONCUR 2014 and TGC 2014. Rome, September 2014.

6.2. Foundations of information hiding

Information hiding refers to the problem of protecting private information while performing certain tasks or interactions, and trying to avoid that an adversary can infer such information. This is one of the main areas of research in Comète; we are exploring several topics, described below.

6.2.1. Additive and multiplicative notions of leakage, and their capacities

Protecting sensitive information from improper disclosure is a fundamental security goal. It is complicated, and difficult to achieve, often because of unavoidable or even unpredictable operating conditions that can lead to breaches in planned security defences. An attractive approach is to frame the goal as a quantitative problem, and then to design methods that measure system vulnerabilities in terms of the amount of information they leak. A consequence is that the precise operating conditions, and assumptions about prior knowledge, can play a crucial role in assessing the severity of any measured vunerability.

In [20] we developed this theme by concentrating on vulnerability measures that are *robust* in the sense of allowing general leakage bounds to be placed on a program, bounds that apply whatever its operating conditions and whatever the prior knowledge might be. In particular we proposed a theory of channel capacity, generalising the Shannon capacity of information theory, that can apply both to additive and to multiplicative forms of a recently-proposed measure known as g-leakage. Further, we explored the computational aspects of calculating these (new) capacities: one of these scenarios can be solved efficiently by expressing it as a Kantorovich distance, but another turns out to be NP-complete.

We also found capacity bounds for arbitrary correlations with data not directly accessed by the channel, as in the scenario of Dalenius's Desideratum.

6.2.2. Compositionality Results for Quantitative Information Flow

In the min-entropy approach to quantitative information flow, the leakage is defined in terms of a minimization problem, which, in case of large systems, can be computationally rather heavy. The same happens for the recently proposed generalization called g-vulnerability. In [28] we studied the case in which the channel associated to the system can be decomposed into simpler channels, which typically happens when the observables consist of several components. Our main contribution was the derivation of bounds on the g-leakage of the whole system in terms of the g-leakages of its components.

6.2.3. LeakWatch: Estimating Information Leakage from Java Programs

Programs that process secret data may inadvertently reveal information about those secrets in their publicly-observable output. In [23] we presented LeakWatch, a quantitative information leakage analysis tool for the Java programming language; it is based on a flexible "point-to-point" information leakage model, where secret and publicly-observable data may occur at any time during a program's execution. LeakWatch repeatedly executes a Java program containing both secret and publicly-observable data and uses robust statistical techniques to provide estimates, with confidence intervals, for min-entropy leakage (using a new theoretical result presented in this paper) and mutual information. We demonstrated how LeakWatch can be used to estimate the size of information leaks in a range of real-world Java programs.

6.2.4. On the information leakage of differentially-private mechanisms

Differential privacy aims at protecting the privacy of participants in statistical databases. Roughly, a mechanism satisfies differential privacy if the presence or value of a single individual in the database does not significantly change the likelihood of obtaining a certain answer to any statistical query posed by a data analyst. Differentially-private mechanisms are often oblivious: first the query is processed on the database to produce a true answer, and then this answer is adequately randomized before being reported to the data analyst. Ideally, a mechanism should minimize leakage—i.e., obfuscate as much as possible the link between reported answers and individuals' data—while maximizing utility—i.e., report answers as similar as possible to the true ones. These two goals are, however, conflicting, and a trade-off between privacy and utility is imposed.

In [13] we used quantitative information flow principles to analyze leakage and utility in oblivious differentially-private mechanisms. We introduced a technique that exploits graph-symmetries of the adjacency relation on databases to derive bounds on the min-entropy leakage of the mechanism. We evaluated utility using identity gain functions, which are closely related to min-entropy leakage, and we derived bounds for it. Finally, given some graph-symmetries, we provided a mechanism that maximizes utility while preserving the required level of differential privacy.

6.2.5. Metric-based approaches for privacy in concurrent systems

In a series of two papers we investigated metric-based techniques for varifying differential privacy in the context of concurrent systems.

The first work [30] was motivated from the one of Tschantz et al., who proposed a verification method based on proving the existence of a stratified family of bijections between states, that can track the privacy leakage, ensuring that it does not exceed a given leakage budget. We improved this technique by investigating state properties which are more permissive and still imply differential privacy. We introduced a new pseudometric, still based on the existence of a family of bijections, but relaxing the relation between them by integrating the notion of amortization, and showed that this results to a more parsimonious use of the privacy budget. We also showed that for the new pseudometric the level of differential privacy is continuous on the distance between the starting states, which makes it suitable for verification.

Continuing this line of work, we studied the pseudometric based on the Kantorovich lifting, which is one of the most popular notions of distance between probabilistic processes proposed in the literature. However, its application in verification is limited to linear properties. In [19], we proposed a generalization which allows to deal with a wider class of properties, such as those used in security and privacy. More precisely, we proposed a family of pseudometrics, parametrized on a notion of distance which depends on the property we want to verify. Furthermore, we showed that the members of this family still characterize bisimilarity in terms of their kernel, and provided a bound on the corresponding distance between trace distributions. Finally, we studied the instance corresponding to differential privacy, and we showed that it has a dual form, easier to compute. We also proved that the typical process-algebra constructs are non-expansive, thus paving the way to a modular approach to verification.

6.2.6. Optimal Geo-Indistinguishable Mechanisms for Location Privacy

With location-based services becoming increasingly more popular, serious concerns are being raised about the potential privacy breaches that the disclosure of location information may induce. In [21] we considered two approaches that have been proposed to limit and control the privacy loss: one is the *geo-indistinguishability* notion developed within Comète, which is inspired by differential privacy, and like the latter it is independent from the side knowledge of the adversary and robust with respect to composition of attacks. The other one is the mechanism of Shokri et al., which offers an optimal trade-off between the loss of quality of service and the privacy protection with respect to a given Bayesian adversary.

We showed that it is possible to combine the advantages of the two approaches: given a minimum threshold for the degree of geo-indistinguishability, we construct a mechanism that offers the maximal utility, as the solution of a linear program. Thanks to the fact that geo-indistinguishability is insensitive to the remapping of a Bayesian adversary, the mechanism so constructed is optimal also in the sense of Shokri et al. Furthermore we proposed a method to reduce the number of constraints of the linear program from cubic to quadratic (with respect to the number of locations), maintaining the privacy guarantees without affecting significantly the utility of the generated mechanism. This lowers considerably the time required to solve the linear program, thus enlarging significantly the size of location sets for which the optimal trade-off mechanisms can still be computed.

6.2.7. A Predictive Differentially-Private Mechanism for Mobility Traces

With the increasing popularity of GPS-enabled handheld devices, location based applications and services have access to accurate and real-time location information, raising serious privacy concerns for their millions of users. Trying to address these issues, the notion of *geo-indistinguishability* was recently introduced, adapting the well-known concept of Differential Privacy to the area of location-based systems. A Laplace-based obfuscation mechanism satisfying this privacy notion works well in the case of a *sporadic* use; Under repeated use, however, *independently* applying noise leads to a quick loss of privacy due to the correlation between the location in the trace.

In [22] we showed that correlations in the trace can be in fact exploited in terms of a *prediction function* that tries to guess the new location based on the previously reported locations. The proposed mechanism tests the quality of the predicted location using a private test; in case of success the prediction is reported otherwise the location is sanitized with new noise. If there is considerable correlation in the input trace, the extra cost of the test is small compared to the savings in budget, leading to a more efficient mechanism.

We evaluated the mechanism in the case of a user accessing a location-based service while moving around in a city. Using a simple prediction function and two budget spending strategies, optimizing either the utility or the budget consumption rate, we showed that the predictive mechanism can offer substantial improvements over the independently applied noise.

6.2.8. A differentially private mechanism of optimal utility for a region of priors

Differential privacy is a notion of privacy that was initially designed for statistical databases, and has been recently extended to a more general class of domains. Both differential privacy and its generalized version can be achieved by adding random noise to the reported data. Thus, privacy is obtained at the cost of reducing the data's accuracy, and therefore their *utility*.

In [31] we considered the problem of identifying *optimal* mechanisms for generalized differential privacy, i.e. mechanisms that maximize the utility for a given level of privacy. The utility usually depends on a prior distribution of the data, and naturally it would be desirable to design mechanisms that are *universally optimal*, i.e., optimal for all priors. However it is already known that such mechanisms do not exist in general. We then characterized maximal *classes of priors* for which a mechanism which is optimal for all the priors of the class *does exist*. We showed that such classes can be defined as convex polytopes in the priors space.

As an application, we considered the problem of privacy that arises when using, for instance, location-based services, and we showed how to define mechanisms that maximize the quality of service while preserving the desired level of geo-indistinguishability.

6.2.9. Compositional analysis of information hiding

Systems concerned with information hiding often use randomization to obfuscate the link between the observables and the information to be protected. The degree of protection provided by a system can be expressed in terms of the probability of error associated to the inference of the secret information. In [14] we considered a probabilistic process calculus to specify such systems, and we studied how the operators affect the probability of error. In particular, we characterized constructs that have the property of not decreasing the degree of protection, and that can therefore be considered safe in the modular construction of these systems. As a case study, we applied these techniques to the Dining Cryptographers, and we derived a generalization of Chaum's strong anonymity result.

6.3. Foundations of Concurrency

Distributed systems have changed substantially in the recent past with the advent of phenomena like social networks and cloud computing. In the previous incarnation of distributed computing the emphasis was on consistency, fault tolerance, resource management and related topics; these were all characterized by *interaction between processes*. Research proceeded along two lines: the algorithmic side which dominated the Principles Of Distributed Computing conferences and the more process algebraic approach epitomized by CONCUR where the emphasis was on developing compositional reasoning principles. What marks the new era of distributed systems is an emphasis on managing access to information to a much greater degree than before.

6.3.1. A Concurrent Pattern Calculus

In [16] we detailed how Concurrent pattern calculus (CPC) drives interaction between processes by comparing data structures, just as sequential pattern calculus drives computation. By generalising from pattern matching to pattern unification, interaction becomes symmetrical, with information flowing in both directions. CPC provides a natural language to express trade where information exchange is pivotal to interaction. The unification allows some patterns to be more discriminating than others; hence, the behavioural theory must take this aspect into account, so that bisimulation becomes subject to compatibility of patterns. Many popular process calculi can be encoded in CPC; this allows for a gain in expressiveness, formalised through encodings.

6.3.2. An Intensional Concurrent Faithful Encoding of Turing Machines

The benchmark for computation is typically given as Turing computability; the ability for a computation to be performed by a Turing Machine. Many languages exploit (indirect) encodings of Turing Machines to demonstrate their ability to support arbitrary computation. However, these encodings are usually by simulating the entire Turing Machine within the language, or by encoding a language that does an encoding or simulation itself. This second category is typical for process calculi that show an encoding of lambda-calculus (often with restrictions) that in turn simulates a Turing Machine. Such approaches lead to indirect encodings of Turing Machines that are complex, unclear, and only weakly equivalent after computation. In [25] we developed an approach to encoding Turing Machines into intensional process calculi that is faithful, reduction preserving, and structurally equivalent. The encoding is demonstrated in a simple asymmetric concurrent pattern calculus before generalised to simplify infinite terms, and to show encodings into Concurrent Pattern Calculus and Psi Calculi.

6.3.3. Expressiveness via Intensionality and Concurrency

Computation can be considered by taking into account two dimensions: extensional versus intensional, and sequential versus concurrent. Traditionally sequential extensional computation can be captured by the lambda-calculus. However, recent work shows that there are more expressive intensional calculi such as SF-calculus. Traditionally process calculi capture computation by encoding the lambda-calculus, such as in the pi-calculus. Following this increased expressiveness via intensionality, other recent work has shown that concurrent pattern calculus is more expressive than pi-calculus. In [26] we formalised the relative expressiveness of all four of these calculi by placing them on a square whose edges are irreversible encodings. This square is representative of a more general result: that expressiveness increases with both intensionality and concurrency.

6.3.4. On the Expressiveness of Intensional Communication

The expressiveness of communication primitives has been explored in a common framework based on the pi-calculus by considering four features: synchronism (asynchronous vs synchronous), arity (monadic vs polyadic data), communication medium (shared dataspaces vs channel-based), and pattern-matching (binding to a name vs testing name equality). In [27] pattern-matching is generalised to account for terms with internal structure such as in recent calculi like Spi calculi, Concurrent Pattern Calculus and Psi calculi. This exploreD intensionality upon terms, in particular communication primitives that can match upon both names and structures. By means of possibility/impossibility of encodings, we showed that intensionality alone can encode synchronism, arity, communication-medium, and pattern-matching, yet no combination of these without intensionality can encode any intensional language.

6.3.5. Weak CCP Bisimilarity with Strong Procedures

Concurrent constraint programming (CCP) is a well-established model for concurrency that singles out the fundamental aspects of asynchronous systems whose agents (or processes) evolve by posting and querying (partial) information in a global medium. Bisimilarity is a standard behavioral equivalence in concurrency theory. However, only recently a well-behaved notion of bisimilarity for CCP, and a CCP partition refinement algorithm for deciding the strong version of this equivalence have been proposed. Weak bisimilarity is a central behavioral equivalence in process calculi and it is obtained from the strong case by taking into account only the actions that are observable in the system. Typically, the standard partition refinement can also be used for deciding weak bisimilarity simply by using Milner's reduction from weak to strong bisimilarity; a technique referred to as saturation. In [17] we demonstrated that, because of its involved labeled transitions, the above-mentioned saturation technique does not work for CCP. We gave an alternative reduction from weak CCP bisimilarity to the strong one that allows us to use the CCP partition refinement algorithm for deciding this equivalence.

6.3.6. Efficient Algorithms for Program Equivalence for Confluent Concurrent Constraint Programming

While the foundations and principles of CCP e.g., semantics, proof systems, axiomatizations, have been thoroughly studied for over the last two decades. In contrast, the development of algorithms and automatic verification procedures for CCP have hitherto been far too little considered. To the best of our knowledge there is only one existing verification algorithm for the standard notion of CCP program (observational) equivalence. In [18] we first showed that this verification algorithm has an exponential-time complexity even for programs from a representative sub-language of CCP; the summation-free fragment (CCP+). We then significantly improved on the complexity of this algorithm by providing two alternative polynomial-time decision procedures for CCP+ program equivalence. Each of these two procedures has an advantage over the other. One has a better time complexity. The other can be easily adapted for the full language of CCP to produce significant state space reductions. The relevance of both procedures derives from the importance of CCP+. This fragment, which has been the subject of many theoretical studies, has strong ties to first-order logic and an elegant denotational semantics, and it can be used to model real-world situations. Its most distinctive feature is that of confluence, a property we exploited to obtain our polynomial procedures.

6.3.7. A Behavioral Congruence for Concurrent Constraint Programming with Nondeterministic Choice

Weak bisimilarity is one of the most representative notions of behavioral equivalence for models of concurrency. As we mentioned earlier, a notion of weak bisimilarity, called weak saturated barbed bisimilarity (wsbb), was recently proposed for CCP. This equivalence improves on previous bisimilarity notions for CCP that were too discriminating and it is a congruence for the choice-free fragment of CCP. In [29], however, we showed that wsbb is not a congruence for CCP with nondeterministic choice. We then introduced a new notion of bisimilarity, called weak full bisimilarity (wfb), and showed that it is a congruence for the full language of CCP. We also showed the adequacy of wfb by establishing that it coincides with the congruence induced by closing wsbb under all contexts. The advantage of the new definition is that, unlike the congruence induced by wsbb, it does not require quantifying over infinitely many contexts.

6.3.8. Abstract Interpretation of Temporal Concurrent Constraint Programs

Timed Concurrent Constraint Programming (tcc) is a declarative model for concurrency offering a logic for specifying reactive systems, i.e. systems that continuously interact with the environment. The universal tcc formalism (utcc) is an extension of tcc with the ability to express mobility. Here mobility is understood as communication of private names as typically done for mobile systems and security protocols. In [15] we considered the denotational semantics for tcc, and we extended it to a "collecting" semantics for utcc based on closure operators over sequences of constraints. Relying on this semantics, we formalized a general framework for data flow analyses of tcc and utcc programs by abstract interpretation techniques. The concrete and abstract semantics we proposed are compositional, thus allowing us to reduce the complexity of data flow analyses. We showed that our method is sound and parametric with respect to the abstract domain. Thus, different analyses can be performed by instantiating the framework. We illustrated how it is possible to reuse abstract domains previously defined for logic programming to perform, for instance, a groundness analysis for tcc programs. We showed the applicability of this analysis in the context of reactive systems. Furthermore, we made use of the abstract semantics to exhibit a secrecy flaw in a security protocol. We also showed how it is possible to make an analysis which may show that tcc programs are suspension free. This can be useful for several purposes, such as for optimizing compilation or for debugging.

6.3.9. Bisimulation for Markov Decision Processes through Families of Functional Expressions

In [24], we transfered a notion of quantitative bisimilarity for labelled Markov processes to Markov decision processes with continuous state spaces. This notion takes the form of a pseudometric on the system states, cast in terms of the equivalence of a family of functional expressions evaluated on those states and interpreted as a real-valued modal logic. Our proof amounted to a slight modification of previous techniques used to prove equivalence with a fixed-point pseudometric on the state-space of a labelled Markov process and making heavy use of the Kantorovich probability metric. Indeed, we again demonstrated equivalence with a fixed-point pseudometric defined on Markov decision processes; what is novel is that we recasted this proof in terms of integral probability metrics defined through the family of functional expressions, shifting emphasis back to properties of such families. The hope is that a judicious choice of family might lead to something more computationally tractable than bisimilarity whilst maintaining its pleasing theoretical guarantees. Moreover, we used a trick from descriptive set theory to extend our results to MDPs with bounded measurable reward functions, dropping a previous continuity constraint on rewards and Markov kernels.

COMMANDS Project-Team

6. New Results

6.1. Highlights of the Year

6.1.1. Optimization of running strategies based on anaerobic energy and variations of velocity

Participant: Frédéric Bonnans.

The paper [10] about running strategies proves Keller’s conjecture. It was highlighted in SIAM Connect,see http://connect.siam.org/insightful-mathematics-for-an-optimal-run/

6.1.2. Research and transfer collaboration in aeronautics with the startup Safety Line

Participants: Frédéric Bonnans, Daphné Giorgi, Stéphan Maindrault, Pierre Martinon.

Following the meeting with the startup Safety Line at Imatch "Optimisation and Control" in october 2013, we conducted a first collaboration of six months on optimizing the fuel consumption of civil airliners. This first step successfully established the proof of concept and was validated by actual test flights in June 2014, leading to a shared patent and the development of a specific module of our software ’Bocop’, included in the tool ’OptiClimb’ developed at Safety Line. Future prospects include improving the numerical robustness of the current tool, as well as expanding the optimization to the cruise flight in addition to the climb phase.

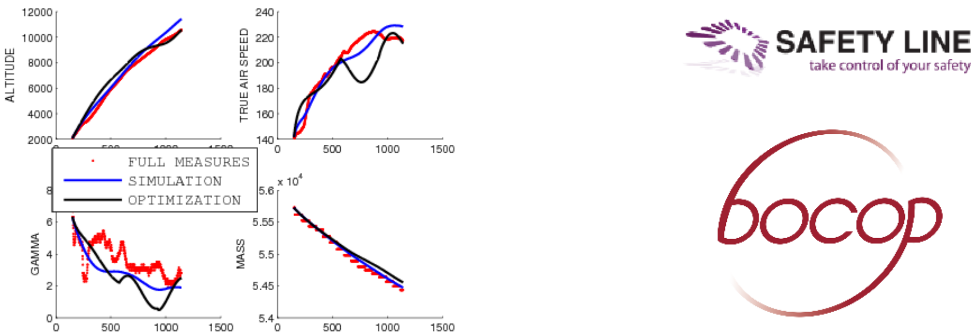


Figure 2. Plane climb phase (Boeing 737)

6.2. Second order analysis of deterministic optimal control problems

Participant: Frédéric Bonnans.

F. Bonnans, with M.S. Aronna (IMPA, Rio de Janeiro) and B.-S. Goh (Curtin U., Miri, Sarawak, Malaysia) obtained in [32] new second order necessary and sufficient optimality conditions for a class of control-affine problems with a scalar control and a scalar state constraint. These optimality conditions extend to the constrained state framework the Goh transform, which is the classical tool for obtaining an extension of the Legendre condition. We propose a shooting algorithm to solve numerically this class of problems and we provide a sufficient condition for its local convergence. We provide examples to illustrate the theory. An article by F. Bonnans, X. Dupuis (Ceremade, U. Dauphine) and L. Pfeiffer (U. Graz) has been published in the SIAM J. Control Optim. on “Second-order necessary conditions in Pontryagin form for optimal control problems” [16].

6.3. Stochastic optimization

6.3.1. Stochastic control

Participant: Frédéric Bonnans.

With J. Gianatti (U. Rosario) and F. Silva (U. Limoges) we obtained an extension of the Sakawa-Shindo algorithm (for computing a solution of the optimality system of a deterministic optimal control problem) to stochastic control problems. The paper is in progress.

6.3.2. Stochastic programming

Participants: Frédéric Bonnans, Nicolas Grebille, Faisal Wahid.

In the framework of the thesis of Nicolas Grébille, we continued our study of decomposition algorithms for a stochastic model of optimal electricity energy production.The energy production is divided in a number of zones. The idea is to constrain the energy flows between these zones, by linear feedback to the demand (wich is a random variable). The coefficients of the feedback are to be optimized. Then the problem is decomposed for each zone (and can then be solved easily by a SDDP type algorithm). We obtained encouraging preliminary numerical results in a three zones problem.
Faisal Wahid developed a mixed integer program model for hydro-power producers participating in the future intra-day French Electricity Balancing Market. He has also formulated the mixed integer stochastic dynamic program model for the more general hydro- bidding under uncertainty. The objective of this model is to produce optimal offer policies in the form of supply curves under a time inhomogeneous Markov process of electricity market clearing prices.

6.3.3. Dynamic programming and error estimates for stochastic control problems with maximum cost

Participants: Athena Picarelli, Hasnaa Zidani.

The paper [14] is concerned with stochastic optimal control for a running maximum cost. A direct approach based on dynamic programming techniques is studied leading to the characterization of the value function as the unique viscosity solution of a second order Hamilton- Jacobi-Bellman (HJB) equation with an oblique derivative boundary condition. A general numerical scheme is proposed and a convergence result is provided. Error estimates are obtained for the semi-Lagrangian scheme. These results can apply to the case of lookback options in finance. Moreover, optimal control problems with maximum cost arise in the characterization of the reachable sets for a system of controlled stochastic differential equations. Some numerical simulations on examples of reachable analysis are included to illustrate our approach.

6.4. Hamilton Jacobi Bellman approach

6.4.1. Optimal feedback control of undamped wave equations by solving a HJB equation

Participants: Hasnaa Zidani, Axel Kröner.

An optimal finite-time horizon feedback control problem for (semi linear) wave equations is studied in [25]. The feedback law can be derived from the dynamic programming principle and requires to solve the evolutionary Hamilton-Jacobi-Bellman (HJB) equation. Classical discretization methods based on finite elements lead to approximated problems governed by ODEs in high dimensional space which makes infeasible the numerical resolution by HJB approach. In the present paper, an approximation based on spectral elements is used to discretize the wave equation. The effect of noise is considered and numerical simulations are presented to show the relevance of the approach

6.4.2. Transmission conditions on interfaces for Hamilton-Jacobi-Bellman equations

Participant: Hasnaa Zidani.

The works [27], [91] deal with deterministic control problems where the dynamic and the running cost can be completely different in two (or more) complementary domains of the space IR^N. As a consequence, the dynamics and running cost present discontinuities at the interfaces of these domains. This leads to a complex interplay that has to be analyzed among transmission conditions to "glue" the propagation of the value function on the interfaces. Several questions arise: how to define properly the value function(s) and what is (are) the right Bellman Equation(s) associated with this problem?. In the case of a simple geometry (namely when the space IR^N is partitioned into two subdomains separated with an interface which is assumed to be a regular hypersurface without any connectedness requirement), [27] discuss different conditions on the hyperplane where the dynamic and the running cost are discontinuous, and the uniqueness properties of the Bellman problem are studied. In this paper we use a dynamical approach, namely instead of working with test functions, the accent is put on invariance properties of an augmented dynamics related to the integrated control system. The comparison principle is accordingly based, rather than on (semi)continuity of the Hamiltonian appearing in the Hamilton–Jacobi–Bellman equation, on some weak separation properties of this dynamics with respect to the stratification.

6.4.3. Control Problems on Stratifiable state-constraints Sets

Participants: Cristopher Hermosilla, Hasnaa Zidani.

This work deals with a state-constrained control problem. It is well known that, unless some compatibility condition between constraints and dynamics holds, the value function has not enough regularity, or can fail to be the unique constrained viscosity solution of a Hamilton-Jacobi-Bellman (HJB) equation. Here, we consider the case of a set of constraints having a strati

ed structure. Under this circumstance, the interior of this set may be empty or disconnected, and the admissible trajectories may have the only option to stay on the boundary without possible approximation in the interior of the constraints. In such situations, the classical pointing quali

cation hypothesis are not relevant. The discontinuous Value Function is then characterized by means of a system of HJB equations on each stratum that composes the state-constraints. This result is obtained under a local controllability assumption which is required only on the strata where some chattering phenomena could occur.

6.4.4. Constrained optimization problems in finite and infinite dimensional spaces

Participant: Cristopher Hermosilla.

We investigate in [39] convex constrained nonlinear optimization problems and optimal control with convex state constraints. For this purpose we endow the interior of constraints set with the structure of Riemannian manifold. In particular, we consider a class of Riemannian metric induced by the squared Hessian of a Legendre functions. We describe in details the geodesic curves on this manifolds and we propose a gradient-like algorithm for constrained optimization based on linear search along geodesics. We also use the Legendre change of coordinates to study the Value Function of a Mayer problem with state constraints. We provide a characterization of the Value Function for this problem as the unique viscosity solution of the Hamilton-Jacobi-Bellman equation.

6.5. Robustness of discontinuous Feedbacks

Participant: Cristopher Hermosilla.

In the paper [40] we study state-constrained discontinuous ordinary differential equations for which the corresponding vector field has a set of singularities that forms a stratification of the state domain. Existence of solutions and robustness with respect to external perturbations of the righthand term are investigated. Moreover, notions of regularity for stratifications are discussed.

6.6. Optimal control of PDEs

6.6.1. Closed-loop optimal control of PDEs

Participant: Axel Kröner.

Stabilization of Burgers equation to nonstationary trajectories A. Kröner and Sérgio S. Rodrigues (RICAM, Linz, Austria) considered in [82] using infinite-dimensional internal controls. Estimates for the dimension of the controller are derived; in the particular case of no constraint in the support of the control a better estimate is derived and the possibility of getting an analogous estimate for the general case is discussed. Numerical examples are presented illustrating the stabilizing effect of the feedback control, and suggesting that the existence of an estimate in the general case analogous to that in the particular one is plausible. In [81] the problem was consided for a finite number of internal piecewise constant controls.

Reduced-order minimum time control of advection-reaction -diffusion systems via dynamic programming Dante Kalise (RICAM, Linz, Austria) and A. Kröner considered in [79]. The authors use balanced truncation for the model reduction part and include a Luenberger observer.

A semi-Lagrangian scheme for L^p-penalized minimum time problems was considered by M. Falcone (Sapienza-Università di Roma, Italy), D. Kalise (RICAM, Austria) and A. Kröner in [78].

6.6.2. Open-loop optimal control of PDEs

Participant: Axel Kröner.

The minimum effort problem for the wave equation K. Kunisch (University of Graz, Austria) and A. Kröner considered in [80]. The problem involves L^∞-control costs which lead to non-differentiability. Uniqueness of the solution of a regularized problem is proven and the convergence of the regularized solutions is analyzed. Further, a semi-smooth Newton method is formulated to solve the regularized problems and its superlinear convergence is shown. Numerical examples confirm the theoretical results.

6.7. Applications in deterministic optimal control

6.7.1. Contrast imaging problem in nuclear magnetic resonance

Participant: Pierre Martinon.

In collaboration with team McTAO (Sophia), we started in 2013 to study the contrast imaging problem in nuclear magnetic resonance, modeled as Mayer problem in optimal control ([58]). Using tools from the Maximum Principle and geometric control, we obtained a first synthesis of locally optimal solutions is given in the single-input case, as well as preliminary results in the bi-input case. This analysis was supported by comprehensive numerical investigations using a combination of indirect shooting (HAMPATH software) and direct method (BOCOP), with a moment-based (LMI) technique to estimate the global optimum.

These results have been extended in 2014, on the theoretical side with the classification of singular extremals ([35]), and on the numerical side with the study of a large number of spins particles subject to spatial inhomogeneities in the magnetic field.

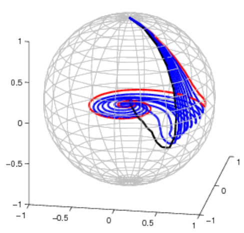


Figure 3. Contrast in quantum control for NMR - Spatial inhomogeneities

6.7.2. Optimal strokes and design for N-link microswimmer

Participant: Pierre Martinon.

Following [71] , we pursued the study of the N-link swimmer, a generalization of the classical Purcell swimmer. We use the model of the Resistive Force Theory to derive the motion equation for the swimmer in a fluid with a low Reynolds number. This allows use to study and solve the optimal swimming problem in the framework of optimal control. We extend our previous study of the optimal strokes by moving to the optimal design of the swimmer. In [72] we provide an estimate of the optimal link ratio for maximal displacement, based on an expansion for small amplitudes. This theoretical result is supported by numerical simulations, that also give some insight on the type of optimal strokes depending on the constraints on the amplitude and deformation speed.

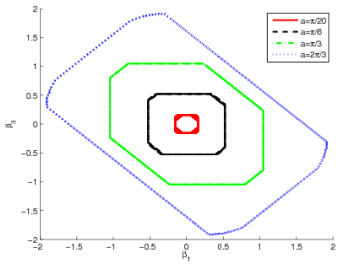


Figure 4. Phase portrait of the optimal stroke w.r.t maximal amplitude

6.7.3. Energy management for a micro-grid

Participants: Frédéric Bonnans, Daphné Giorgi, Benjamin Heymann, Stéphan Maindrault, Pierre Martinon.

We study the energy management problem for a microgrid including a diesel generator and a photovoltaic plant with a battery storage system. The objective is to minimize the total operational cost over a certain timeframe, primarily the diesel consumption, while satisfying a prescribed power load. After reformulation, the decision variables can be reduced to the charging/discharging power for the battery system. We take into account the switching cost for the diesel generator, the non-convex objective, and the long-term aging of the batteries. We solve this problem using a continuous optimal control framework, with both a direct transcription method (time discretization) and a Dynamic Programming method (Hamilton Jacobi Bellman). This project is a collaboration between team COMMANDS (Inria Saclay, France) and Centro de Energia (Universidad de Chile, Chile). A first paper is currently in preparation, while ongoing studies include comparison with the existing MILP approach, more refined battery aging models, and modeling the stochastic nature of the photovoltaic power and power load.

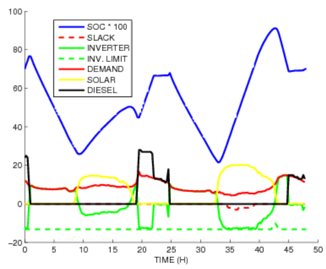


Figure 5. Microgrid management - Winter day sample case

COMPSYS Project-Team

6. New Results

6.1. Highlights of the Year

For 2014, from the point of view of organization, funding, collaborations, the main points to highlight are:

- Christophe Alias and Alexandru Plesco have co-founded the XTREMLOGIC start-up in January 2014 (see Section 7.2), following the incubation of Zettice. XTREMLOGIC recently won the "concours région rhône-alpes" grant in November 2014 (40k).
- Tomofumi Yuki was hired as an Inria researcher and became a permanent member of Compsys.
- The 1988 "Array Expansion" seminal paper of Paul Feautrier has been selected for the 25th Anniversary Volume of the ACM International Conference on Supercomputing (ICS) together with 34 other papers selected from the 1800 papers published from 1987 to 2011. A short "reminescence" paper [13] was written for the occasion.
- The team was evaluated in Nov. 2014 by the HCERES (new name of AERES), as part of the LIP lab evaluation. The report has not been received yet.

From a scientific point of view, the shift, in Compsys III, towards the analysis of parallel programs and the extensions of the polyhedral model, both in terms of techniques and applications, is continuing, see the section "New Results", in particular:

- The design (by Christophe Alias and Alexandru Plesco) of a HLS compiler technology (see Section 6.2), patented by Inria [12] and transferred to XTREMLOGIC under an Inria licence (see Section 5.5).
- Two new static analyses: a more precise array bound check analysis [9] (see Section 6.3) and a more scalable termination algorithm for C programs (see Section 6.4).
- A novel equivalence-checking algorithm [7] modulo associativity/commutativity, which is a first step towards semantic program transformations (see Section 6.5).
- A groundbreaking introduction of polyhedral techniques for the analysis of parallel programs, in particular X10 (see [29] and [6]) and OpenStream (see Section 6.6).
- A seminal paper [5] introducing polynomial techniques in program analysis and compilation (see Section 6.7).
- Innovative contributions on parametric tiling [8], [3], [4] as extensions of the polyhedral model (see Sections 6.8 and 6.9).

6.2. Data-Aware Process Networks

Participants: Christophe Alias, Alexandru Plesco [XTREMLOGIC start-up].

Process networks are execution models expressing naturally the parallelism of a computation. They are a natural intermediate representation for high-level synthesis tools, where the front-end extracts the parallelism and produces a process network and the back-end compiles the process network to the target architecture.

In that context, we have defined a new model of process network that fits HLS-specific constraints, the data-aware process network (DPN). Our model makes explicit the communications with the central memory and the parallel access to channels, and is close enough to the hardware constraints to be translated directly to a circuit. We show how to compile an imperative program to a DPN, so as to optimize both the I/O and the parallelism, while using the polyhedral model.

DPNs are used as the intermediate representation for the HLS compiler suite of the XTREMLOGIC start-up. They are generated from C programs by the Dcc compiler (see Section 5.5). The apparatus underlying the DPN synchronizations has been patented by Inria [12].

6.3. Preventing from Out-of-Bound Memory Accesses

Participants: Laure Gonnord, Fernando Pereira [Univ. Mineas Gerais, Brasil].

The C programming language does not prevent out-of-bounds memory accesses. There exist several techniques to secure C programs; however, these methods tend to slow down these programs substantially, because they populate the binary code with runtime checks. To deal with this problem, we designed and tested two static analyses (symbolic region and range analysis), which we combine to remove the majority of these guards.

In addition to the analyses themselves, we brought two other contributions:

- First, we described live-range splitting strategies that improve the efficiency and the precision of our analyses.
- Secondly, we showed how to deal with integer overflows, a phenomenon that can compromise the correctness of static algorithms validating memory accesses.

We validated our claims by incorporating our findings into AddressSanitizer (see https://code.google.com/p/address-sanitizer/). We generated SPEC CINT 2006 code that is 17% faster and 9% more energy efficient than the code originally produced by this tool. Furthermore, our approach is 50% more effective than Pentagons, a state-of-the-art analysis to sanitize memory accesses. This work was published at the OOPSLA 2014 conference [9].

6.4. Scaling Termination Proofs

Participants: Laure Gonnord, Gabriel Radanne [ENS Rennes], David Monniaux [CNRS/VERIMAG], Fernando Pereira [Univ. Mineas Gerais, Brasil], Raphael Rodrigues [Univ. Mineas Gerais, Brasil].

In [15], we presented a new algorithm adapted from scheduling techniques to synthesize (multi-dimensional) affine functions from general flowcharts programs. But, as for other methods, our algorithm tried to solve linear constraints on each control point and each transition, which can lead to quasi-intractable linear programming instances. In contrast to these approaches, we proposed a new algorithm based on the following observations:

- Searching for ranking functions for loop headers is sufficient to prove termination.
- Furthermore, there exist loops such that there is a linear lexicographic ranking function that decreases along each path inside the loop, from one loop iteration to the next, but such that there is no lexicographic linear ranking function that decreases at each step along these paths. For these reasons, it is tempting to treat each path inside a loop as a single transition.

Unfortunately the number of paths may be exponential in the size of the program, thus the constraint system may become very large, even though it features fewer variables. To face this theoretical complexity, even though the number of paths may be large, we argue that, in practice, few of them actually matter in the constraint system (we formalize this concept by giving a characterization as geometric extremal points). Our algorithm therefore builds the constraint system lazily, taking paths into account *on demand*.

In 2014, we consolidated this approach with a work on complexity issues (inspired by [19]) and a new implementation: Termite (see Section 5.13). A corresponding paper is currently under submission for PLDI.

With Fernando Pereira's group in Brazil, we also studied the relevance of fast and simple solutions to compute approximations of the number of iterations of loops (*loop trip count*) of imperative real-world programs. The context of this work is the use of these approximations in compiler optimizations: most of the time, the optimizations yield greater benefits for large trip counts, and are either innocuous or detrimental for small ones. In our paper published at WST'14 [10], we have shown that, most of the time, there is no need to use computationally-expensive state-of-the-art methods to compute (an approximation of) it. We support our position with an actual case study. We show that a fast predictor can be used to speedup the JavaScript JIT compiler of Firefox - one of the most well-engineered runtime environments in use today.

6.5. Equivalence-Checking of Programs with Reductions

Participants: Guillaume Iooss, Christophe Alias, Sanjay Rajopadhye [Colorado State University, USA].

Program equivalence is a well-known problem with a wide range of applications, such as algorithm recognition, program verification, and program optimization. This problem is also known to be undecidable if the class of programs is rich enough, in which case semi-algorithms are commonly used.

We focused on programs represented as systems of affine recurrence equations (SARE), defined over parametric polyhedral domains, a well-known formalism for the *polyhedral model*. SAREs include, as a proper subset, the class of affine control loop programs. Several semi-algorithms for program equivalence were already proposed for this class. Some take into account algebraic properties such as associativity and commutativity. To the best of our knowledge, none of them manage reductions, i.e., accumulations of a *parametric* number of sub-expressions using an associative and commutative operator. Our main contribution has been a new semi-algorithm to manage reductions. In particular, we outlined the ties between this problem and the perfect matching problem in a parametric bipartite graph.

This work was published at the SAS 2014 conference [7].

6.6. Analysis and Transformation of Parallel Programs

Participants: Albert Cohen [Inria/PARKAS], Alain Darte, Paul Feautrier, Abdoulaye Gamatie [CNRS/LIRMM], Laure Gonnord, Alain Ketterlin [Inria/CAMUS], Sanjay Rajopadhye [Colorado State University], Vijay Saraswat [IBM Research], Eric Violard [Inria/CAMUS], Tomofumi Yuki.

While, historically, Compsys has applied polyhedral analysis to sequential programs, it was recently realized that it also applies to parallel programs, with the aim of checking their correctness or improving their performance. The prospect of having to program exascale architectures, with their millions of cores, has led to the development of new programming languages, whose objective is to increase the programmer productivity. Compsys has applied polyhedral techniques to synchronous languages (see [25], [26] and previous activity reports), to IBM's high-productivity language X10, and, in the context of the ManycoreLabs project, to a streaming language, OpenStream, developed by Albert Cohen's group.

X10 is based on the creation of independent *activities* (light-weight threads), which can synchronize either by a generalization of the fork/join scheme, or with *clocks*, an improved version of the familiar barriers. X10 is deadlock-free by construction but it is the programmer responsibility to insure determinism by a proper use of synchronizations. Non-determinism bugs may have a very low occurrence probability thus be very difficult to detect by testing, hence the interest for detecting races at compile time. In collaboration with CSU (S. Rajopadhye, T. Yuki) and IBM (V. Saraswat), we extended array dataflow analysis to polyhedral clock-free X10 programs [29]. We have been working on clocked programs too: race detection becomes undecidable [30], but realistic problems may still be solved by heuristics.

As a side-effect of this work, we have shown in cooperation with Eric Violard and Alain Ketterlin (Inria Team Camus, Strasbourg) that clocks can be removed and replaced by `async/finish` constructs without modifying the program semantics [6]. While this transformation incurs a large overhead for general programs, in the polyhedral case the overhead is negligible, thus improving the program performance.

In contrast to X10, OpenStream is deterministic by construction, but may have deadlocks. A usual way of disproving deadlocks is by exhibiting a schedule for the program operations, a well-known problem for polyhedral programs, where dependences can be described by affine constraints. In the case of OpenStream, communications use one-dimensional channels and, in a form of linearization, give rise to polynomial dependences for polyhedral OpenStream codes. In a ManycoreLabs project deliverable (see Section 7.1), we have formalized the problem and proved that deadlock detection is undecidable in general.

6.7. Handling Polynomials for Program Analysis and Transformation

Participant: Paul Feautrier.

As shown in the previous section, many problems in parallel programs analysis and verification can be reduced to proving or disproving properties of polynomials in the variables of the program. For instance, so-called “linearizations” (replacing a multi-dimensional object by a one-dimensional one) generate polynomial access functions. These polynomials then reappear in dependence testing, scheduling, and invariant construction. This is also the case in OpenStream where nested loops act on one-dimensional streams. What is needed here is a replacement for the familiar emptiness tests and for Farkas lemma (deciding whether an affine form is positive inside a polyhedron).

Recent mathematical results by Handelman and Schweighofer on the *Positivstellensatz* allow one to devise algorithms that are able to solve these problems. The difference is that one gets only sufficient conditions, and that complexity is much higher than in the affine cases. A paper presenting applications of these ideas to three use cases – dependence testing, scheduling, and transitive closure approximation – will be presented at the 5th International Workshop on Polyhedral Compilation Techniques (IMPACT'15) [5] in Amsterdam in January 2015.

6.8. Parametric Loop Tiling with Constant Aspect Ratio

Participants: Guillaume Iooss, Christophe Alias, Sanjay Rajopadhye [Colorado State University, USA].

Parametric tiling is a well-known transformation which is widely used to improve locality, parallelism, and granularity (see also the next section for more details). However, parametric tiling is also a non-linear transformation and this prevents polyhedral analysis or further polyhedral transformation after parametric tiling. It is therefore generally applied during the code generation phase.

To address this issue, we presented a method to remain in a polyhedral representation, in a special case of parametric tiling where all the dimensions are tiled and all tile sizes are constant multiples of a single tile size parameter. We call this *Constant Aspect Ratio Tiling*. We showed how to mathematically transform a polyhedron and an affine function into their tiled counterpart, which are the two main operations needed in such a transformation.

The approach is now implemented, and has been tested successfully on several kernels commonly used in the community (matrix multiply, jacobi 1D, jacobi 2D). A corresponding paper was published at the IMPACT 2014 workshop [8].

6.9. Exact and Approximated Data-Reuse Optimizations for Tiling with Parametric Sizes

Participants: Alain Darte, Alexandre Isoard.

Loop tiling is a loop transformation widely used to improve spatial and temporal data locality, to increase computation granularity, and to enable blocking algorithms, which are particularly useful when offloading kernels on computing units with smaller memories. When caches are not available or used, data transfers and local storage must be software-managed, and some useless remote communications can be avoided by exploiting data reuse between tiles. An important parameter of tiling is the sizes of the tiles, which impact the size of the required local memory. However, for most analyses involving several tiles, which is the case for inter-tile data reuse, the tile sizes induce non-linear constraints, unless they are numerical constants. This complicates or prevents a parametric analysis with polyhedral optimization techniques.

We showed that, when tiles are executed in sequence along tile axes, the parametric (with respect to tile sizes) analysis for inter-tile data reuse is nevertheless possible, i.e., one can determine, at compile-time and in a parametric fashion, the copy-in and copy-out data sets for all tiles, with inter-tile reuse, as well as sizes for the induced local memories. When approximations of transfers are performed, the situation is much more complex, and involves a careful analysis to guarantee correctness when data are both read and written. We provide the mathematical foundations to make such approximations possible, thanks to the introduction of the concept of *pointwise functions*. Combined with hierarchical tiling, this result opens perspectives for the automatic generation of blocking algorithms, guided by parametric cost models, where blocks can be pipelined and/or can contain parallelism. Previous work on FPGAs and GPUs already showed the interest and feasibility of such automation with tiling, but in a non-parametric fashion. Our method is currently implemented with the `iscc` calculator of `ISL`, a library for the manipulation of integer sets defined with Presburger arithmetic, a complete implementation within the PPCG compiler is in progress.

We believe that our approximation technique can be used for other applications linked to the extension of the polyhedral model as it turns out to be fairly powerful. Our future work will be to derive efficient approximation techniques, either because the program cannot be fully analyzable, or because approximations can speed-up or simplify the results of the analysis without losing much in terms of memory transfers and/or memory sizes.

A preliminary version of this work has been presented at the IMPACT'14 workshop [3]. A revised version has been accepted for publication at the International Conference on Compiler Construction (CC'15) [4].

6.10. Studying Optimal Spilling in the Light of SSA

Participants: Florian Brandner [ENSTA ParisTech], Quentin Colombet [Apple, Cupertino], Alain Darte.

Recent developments in register allocation, mostly linked to static single assignment (SSA) form, have shown the benefits of decoupling the problem in two phases: a first spilling phase places load and store instructions so that the register pressure at all program points is small enough, a second assignment and coalescing phase maps the variables to physical registers and reduces the number of move instructions among registers. At the end of Quentin Colombet's PhD thesis, we focused on the first phase, for which many open questions remained: in particular, we studied the notion of optimal spilling (what can be expressed?) and the impact of SSA form (does it help?). To identify the important features for optimal spilling on load-store architectures, we developed a new integer linear programming formulation, more accurate and expressive than previous approaches. Among other features, we can express SSA ϕ-functions, memory-to-memory copies, and the fact that a value can be stored simultaneously in a register and in memory. Based on this formulation, we presented a thorough analysis of the results obtained for the SPECINT 2000 and EEMBC 1.1 benchmarks, from which we drawed the following conclusions: a) rematerialization is extremely important, b) SSA complicates the formulation of optimal spilling, especially because of memory coalescing when the code is not in conventional SSA (CSSA), c) micro-architectural features are significant and thus have to be accounted for, which is not the case with standard cost functions, d) significant savings can be obtained in terms of static spill costs, cache miss rates, and dynamic instruction counts, e) however, cost models based only on static spill costs are not always relevant, in particular when spilling is "at the limit": in this situation, bad interactions with register coalescing and post-pass scheduling can be exacerbated and it may be better to spill a bit more. This important observation indicates that more research is needed to explore alternative cost models that reliably guide spilling.

Parts of this work were already published at CASES 2011. The publication at ACM Transactions on Architecture and Code Optimization [1] contains more detailed discussions, more examples illustrating new concepts and existing approaches, and additional experiments covering the observed worst-case behavior, a new post-latency heuristic, and empiric evidence showing why static spill costs are a poor metric. Three configurations were added: Appel and George under SSA, Koes and Goldstein, and the heuristic of Braun and Hack. This work was partly supported by the Mediacom contract with STMicroelectronics (ended in 2013).

CONVECS Project-Team

6. New Results

6.1. New Formal Languages and their Implementations

LNT is a next generation formal description language for asynchronous concurrent systems, which attempts to combine the best features of imperative programming languages and value-passing process algebras. LNT is increasingly used by CONVECS for industrial case studies and applications (see § 6.5) and serves also in university courses on concurrency, in particular at ENSIMAG (Grenoble) and at Saarland University.

6.1.1. Translation from LNT to LOTOS

Participants: Hubert Garavel, Frédéric Lang, Wendelin Serwe.

In 2014, the translator from LNT to LOTOS was further improved. In addition to bug fixes and removal of incorrect warnings emitted by the translator itself or by the C compiler on the generated code, the following enhancements have been brought: the LNT language was extended with a "`!representedby`" pragma for processes, and a "`only if`" statement to concisely express guarded commands; the translator now performs static analysis and warns about unused variables, unused "`in`" or "`in out`" parameters, useless (deterministic or nondeterministic) assignments to variables, "`in out`" parameters that are never assigned, and dubious synchronizations between processes; checks for underflow/overflow on natural and integer numbers are now activated by default. The translator also generates better LOTOS code, and the LNT reference manual was shortened and updated in many places.

6.1.2. Translation from LOTOS to Petri nets and C

Participants: Hubert Garavel, Wendelin Serwe.

The LOTOS compilers CAESAR and CAESAR.ADT, which were once the flagship of CADP, now play a more discrete role since LNT (rather than LOTOS) has become the recommended specification language of CADP. Thus, CAESAR and CAESAR.ADT are mostly used as back-end translators for LOTOS programs automatically generated from LNT or other formalisms such as Fiacre, and are only modified when this appears to be strictly necessary.

In 2014, the CAESAR compiler has been modified to tolerate LOTOS specifications that would be normally rejected under the ISO/IEC 8807 standard definition of LOTOS. The first change extends the visibility scope of local definitions when the global definitions are empty. The second change uses the type information of process definitions to better resolve overloading ambiguities in expressions passed as actual parameters to process calls.

Conversely, CAESAR was made stricter by rejecting at compile-time LOTOS specifications containing out-of-bound constants, even if such constants are never used.

Performance has been increased by adding or strengthening a number of optimizations concerning, e.g., internal data structures, Boolean guards that can be statically evaluated, values belonging to singleton sorts, disconnected or otherwise unreachable Petri net places and transitions, etc.

The CAESAR.BDD tool of CADP, which analyzes hierarchical Petri nets generated from higher-level specifications (e.g., LOTOS or LNT) has been significantly enhanced. The semantic model accepted by CAESAR.BDD has been made more general and given the new name of NUPN (*Nested-Units Petri Nets*). The definition and theoretical properties of NUPN have been formalized.

The textual syntax for NUPN has been extended with pragmas intended to retain useful properties of non-ordinary and/or non-safe Petri nets translated to NUPN. An XML syntax for NUPN (compatible with the ISO standard PNML for the representation of Petri nets) has been defined and adopted by the Model Checking Contest [0]. A translator from PNML to NUPN has been developed at LIP6 (Paris, France).

[0] http://mcc.lip6.fr/nupn.php

The CAESAR.BDD tool has been updated accordingly, and extended to perform stricter checks and compute more structural and behavioral properties of NUPN models. CAESAR.BDD has been intensively used to correct the descriptions of the Model Checking Contest benchmarks: a first campaign (January-February 2014) detected 9 errors in structural properties and 8 errors in behavioral properties, and a second campaign (April 2014) revealed 23 more errors. CAESAR.BDD has also been used to automatically generate new benchmarks, together with their descriptions.

6.1.3. Translation from GRL to LNT

Participants: Fatma Jebali, Frédéric Lang, Eric Léo, Radu Mateescu.

In the context of the Bluesky project (see § 8.1.2.1), we study the formal modeling of GALS (*Globally Asynchronous, Locally Synchronous*) systems, which are composed of several synchronous subsystems evolving cyclically, each at its own pace, and communicating with each other asynchronously. Designing GALS systems is challenging due to both the high level of (synchronous and asynchronous) concurrency and the heterogeneity of computations (deterministic and nondeterministic). To bring our formal verification techniques and tools closer to the GALS paradigm, we designed a new formal language named GRL (*GALS Representation Language*), as an intermediate format between GALS models and purely asynchronous concurrent models. GRL combines the main features of synchronous dataflow programming and asynchronous process calculi into one unified language, while keeping the syntax homogeneous for better acceptance by industrial GALS designers. GRL allows a modular composition of synchronous systems (blocks), environmental constraints (environments), and asynchronous communication mechanisms (mediums), to be described at a level of abstraction that is appropriate to verification. GRL also supports external C and LNT code.

In 2014, we have continued to enhance the syntax and the formal semantics of GRL. We have written a detailed research report (82 pages) [25] containing the complete definition of the syntax, static semantics, and dynamic semantics (in the form of structural operational semantics rules), and also illustrating the checking of dynamic semantics rules on several examples of GRL programs. A paper describing GRL has been published in an international conference [14].

To equip GRL with verification features, we formally defined a translation from GRL to LNT. GRL blocks are translated into LNT functions, possibly encapsulated within LNT wrapper processes to enable asynchronous communication, whereas GRL environments and mediums are directly translated into LNT processes. The asynchronous composition of blocks, environments, and mediums is translated to an LNT parallel composition of the corresponding processes.

Using the SYNTAX and LOTOS NT compiler construction technology [44], we have developed a translator named GRL2LNT (about $25,000$ lines of code), allowing an LNT program to be generated from a GRL specification automatically. GRL2LNT performs the lexical and syntactic analysis of GRL programs, together with almost all static semantic checks specified in its formal definition [25]. A stable version of GRL2LNT has been released in 2014. Additionally, we have developed an OPEN/CAESAR-compliant compiler GRL.OPEN (based on GRL2LNT and LNT.OPEN), which makes possible the on-the-fly exploration of the LTS underlying a GRL specification using CADP. We have also built a test base containing about 250 (correct and incorrect) GRL programs, and used it for non-regression testing of GRL2LNT. The correct GRL programs represent about $7,000$ lines of code and produce about $18,000$ lines of LNT code after translation using GRL2LNT.

A paper describing the formal verification of GALS systems using GRL and CADP, with a focus on the translation from GRL to LNT, has been submitted to an international conference [28].

6.1.4. Coverage Analysis for LNT

Participants: Gwen Salaün, Lina Ye.

In the classic verification setting, the designer has a specification of a system in a value-passing process algebra, a set of temporal properties to be verified on the corresponding LTS model, and a data set of examples (test cases) for validation purposes. At this stage, building the set of validation examples and debugging the specification is a complicated task, in particular for non-experts.

We propose a new framework for debugging value-passing process algebra through coverage analysis and we illustrate our approach with LNT. We define several coverage notions before showing how to instrument the specification without affecting original behaviors. Our approach helps one to improve the quality of a data set of examples used for validation purposes, but also to find ill-formed decisions, dead code, and other errors in the specification. We have implemented a tool for automating our debugging approach, and applied it to several real-world case studies in different application areas.

In 2014, a paper has been accepted in an international conference [19].

6.1.5. Other Language Developments

Participants: Hugues Evrard, Hubert Garavel, Frédéric Lang, Eric Léo, Wendelin Serwe.

The ability to compile and verify formal specifications with complex, user-defined operations and data structures is a key feature of the CADP toolbox since its very origins. A long-run effort has been recently undertaken to ensure a uniform treatment of types, values, and functions across all the various CADP tools.

In 2014, convergence between the LOTOS, LNT, BCG, and XTL data-type libraries has been increased by defining common libraries for eight predefined types: Boolean, Natural, Integer, Real, Character, String, Raw, and Gate. These libraries gather in the same place definitions of types, constants, and functions that were previously disseminated across different tools. Additionally, systematic checks for underflows and overflows have been set for the Natural and Integer types. Also, unprintable characters and C-like escape sequences are now uniformly handled by the Character, String, and Raw types.

To support the LNT language in the Emacs/XEmacs, jEdit, and Vim editors, configuration files have been added or updated, which provide for syntax highlighting/coloring, and enable autocompletion in Emacs using YASnippet.

6.2. Parallel and Distributed Verification

6.2.1. Distributed Code Generation for LNT

Participants: Hugues Evrard, Frédéric Lang.

Rigorous development and prototyping of a distributed verification algorithm in LNT involves the automatic generation of a distributed implementation. For the latter, a protocol realizing process synchronization is required. As far as possible, this protocol must itself be distributed, so as to avoid the bottleneck that would inevitably arise if a unique process would have to manage all synchronizations in the system. A particularity of such a protocol is its ability to support *branching synchronizations*, corresponding to situations where a process may offer a choice of synchronizing actions (which themselves may nondeterministically involve several sets of synchronizing processes) instead of a single one. Therefore, a classical barrier protocol is not sufficient and a more elaborate synchronization protocol is needed.

Using a synchronization protocol that we verified formally in 2013, we developed a prototype distributed code generator, named DLC (*Distributed LNT Compiler*), which takes as input the model of a distributed system described as a parallel composition of LNT processes.

In 2014, we continued the development of DLC. We improved the performances of DLC generated code by reducing the number of protocol messages when one or several processes are ready on a single gate. We experimented this optimization on a set of processes running on different computers and synchronizing all together on a single barrier interaction (i.e., all processes are ready on a single gate). In this situation, DLC now generates code that is faster than Java or Erlang.

The distributed program generated by DLC would be of little interest if it could not interact with its environment (e.g., users through human-computer interfaces, or other systems, such as databases, Web services, etc.). Therefore, we designed a mechanism to embed user-defined C functions, called *hook functions*, into the code generated by DLC. Hook functions are triggered on events related to actions in the system. This allows system actions to be, e.g., monitored by the user or controlled by external conditions. Using hook functions, the code generated by DLC can thus both take an account of and have an effect on its environment.

In order to demonstrate DLC on a real-world example, we applied it to the recent Raft [0] consensus algorithm [60]. We wrote an LNT specification of a simple key-value store made fault tolerant by replication of commands using the Raft consensus algorithm. During the modeling phase, we found a missing transition in the TLA+ specification of the protocol. We signaled it to the authors [0], who corrected the TLA+ specification. We used hook functions to implement interaction with the replicated store from external clients. The distributed implementation generated by DLC was successfully tested on clusters of the Grid5000 platform. We presented an overview of DLC, the hook functions and the Raft experiment in an article that has been accepted for publication in an international conference [12].

6.3. Timed, Probabilistic, and Stochastic Extensions

6.3.1. Model Checking for Extended PCTL

Participants: Hubert Garavel, Radu Mateescu, Jose Ignacio Requeno.

In the context of the SENSATION project (see § 8.2.1.1), we study the specification and verification of quantitative properties of concurrent systems.

In 2014, we defined an extension of PCTL (*Probabilistic Computation Tree Logic*) [49] with the manipulation of data values and actions. This logic is interpreted on extended DTMCs (*Discrete-Time Markov Chains*) containing visible transitions, labeled with channel names and data values, in addition to probabilistic transitions. Extended PCTL makes possible the specification of temporal properties involving discrete time, probabilities, and data values.

We devised a prototype model checker for extended PCTL in the form of an XTL library describing the denotational semantics of all PCTL operators (both primitive and derived ones), accompanied by external C code implementing the algorithms for LTS exploration and numerical computation of probabilities. The high-level programming language constructs of XTL (iterators, sets in comprehension, parameterized macro-definitions) allowed us to easily implement the advanced features (filters on arithmetic and logical operators, computation of probabilities, experiments over data series, etc.) of established probabilistic model checkers, such as PRISM [54]. Also, the manipulation of data values in XTL allows one to specify properties in which probabilities and discrete time deadlines depend on the values of state variables, a feature currently not provided by PRISM.

To experiment and cross-check our extended PCTL library w.r.t. PRISM, we developed an automated translator from the (state-based) DTMCs used by PRISM into the (action-based) DTMCs in BCG format used by CADP. State information is represented by means of special self-looping transitions containing the values of state variables, which are properly handled during the evaluation of probabilistic temporal operators.

The experiments we performed with our extended PCTL library on various examples of DTMCs (produced from communication protocols, chemical reactions, hazard games, etc.) showed a performance comparable to (explicit-state) PRISM for pure PCTL formulas.

Furthermore, in addition to many bug fixes, the XTL compiler and its XTL_EXPAND preprocessor have been strengthened to better detect and report potential mistakes in source XTL specifications. In particular, vacuity checks have been introduced, which warn the user when no label in a BCG graph has the right number of fields or the appropriate field types to satisfy an XTL label match expression (previously, this expression would silently evaluate to false).

The type checking system of XTL and its list of predefined functions have been extended to support the new Natural and Raw types of the BCG format, and to properly distinguish between Natural and Integer values, and Raw and String values, while achieving a high degree of backward compatibility. In particular, XTL now uses type information from the BCG labels to better solve overloading in label offers, so that certain XTL programs that were formerly invalid are now accepted. Finally, it is now possible to use the predefined types and functions of XTL when defining temporal operators directly using external C code.

[0] http://raftconsensus.github.io
[0] https://groups.google.com/forum/#!topic/raft-dev/yu-wOUx-gnA

6.4. Component-Based Architectures for On-the-Fly Verification

6.4.1. Property-Dependent Reductions for the Modal Mu-Calculus

Participant: Radu Mateescu.

In collaboration with Anton Wijs (Technical University of Eindhoven), we proposed a new method for enhancing the performance of model checking a temporal formula on an LTS by reducing the LTS as much as possible depending on the formula prior to (or simultaneously with) the verification. Given an LTS and a formula, the method consists of two steps:

- The maximal set of actions that one can hide (i.e., rename into the internal action τ) in the LTS without disturbing the interpretation of the formula is computed, and those actions are hidden in the LTS. This works for any formula of the full modal μ-calculus (i.e., of arbitrary alternation depth) and provides the highest potential for reducing the LTS, and hence for improving verification performance, w.r.t. that formula.
- The LTS is reduced modulo an equivalence relation preserving the formula. The reduction can be done before verification, either by constructing the LTS explicitly and using the direct minimization features provided by the BCG_MIN tool, or by constructing the minimized LTS incrementally using the compositional verification features provided by EXP.OPEN and SVL. The reduction can be also done simultaneously during verification, by detecting τ-confluent transitions and prioritizing them against their neighbors.

We defined a μ-calculus fragment, named Lμ-dsbr, and shown its adequacy w.r.t. divergence-sensitive branching bisimulation (divbranching for short). We also shown that Lμ-dsbr is equally expressive to the μ-ACTL$\setminus$X logic, an extension of ACTL$\setminus$X (Action-based CTL without the next time operator) with fixed point operators [39], [40]. This result also implies the adequacy w.r.t. divbranching of μ-ACTL$\setminus$X, which was previously shown to be adequate w.r.t. strong bisimulation.

We experimented our method using the EVALUATOR model checker on various examples of protocols and distributed systems, by specifying the temporal properties in MCL and reducing the LTSs modulo strong and divbranching bisimulation. The experiments showed performance enhancements both in execution time (reduction by a factor 4 for strong bisimulation and 20 for divbranching) and memory consumption (reduction by a factor 2 for strong bisimulation and 5 for divbranching).

We also built a prototype MCL library regrouping the temporal operators of ACTL$\setminus$X (which were already present in CADP) and the modal and temporal operators of Lμ-dsbr (which were newly added). Used in conjunction with the Boolean and fixed point operators of MCL, the operators of this library can be used to specify temporal formulas adequate w.r.t. divbranching, which allows one to reduce the LTS modulo this equivalence (after applying maximal hiding) and to increase the performance of verification accordingly. An article has been published in an international journal [8].

6.4.2. Compositional Verification

Participants: Hubert Garavel, Frédéric Lang.

The CADP toolbox contains various tools dedicated to compositional verification, among which EXP.OPEN, BCG_MIN, BCG_CMP, and SVL play a central role. EXP.OPEN explores on the fly the graph corresponding to a network of communicating automata (represented as a set of BCG files). BCG_MIN and BCG_CMP respectively minimize and compare behavior graphs modulo strong or branching bisimulation and their stochastic extensions. SVL (*Script Verification Language*) is both a high-level language for expressing complex verification scenarios and a compiler dedicated to this language.

In 2014, we corrected 2 bugs in EXP.OPEN, 6 bugs in BCG_MIN and BCG_CMP, and 5 bugs in SVL. We also enhanced these tools as follows:

- We corrected the diagnostic generation algorithm of BCG_CMP, which sometimes generated irrelevant diagnostics.

- We improved the messages displayed by SVL and EXP.OPEN when generating an LTS from a composition expression using the *smart reduction* strategy [38], so that the user can follow more easily the selected composition order.
- Following the recent progress made on the development of the language LNT (see § 6.1), the syntax of the SVL and EXP languages for comments, gate typing, and the "`par`", "`hide`", "`rename`", "`cut`", and "`prio`" operators was extended to be compatible with the syntax of LNT. This enables composition expressions (including comments, channel typing, etc.) copied from LNT programs to be pasted in SVL scripts while requiring as few syntactic changes as possible.
- The "`verify`" operator has been generalized to give access to all three model checkers of CADP (EVALUATOR 3, EVALUATOR 4, and XTL). A new statement "`|=`" has been added to SVL, which enables MCL and XTL formulas to be directly written in an SVL script, rather than being stored in external files.
- To provide for requirements expression and traceability in SVL, we introduced two new statements, "`property`" and "`check`", which increase the readability and good structure of SVL scripts by allowing to define and verify properties, each of which is given a name, instantiable parameters, an informal textual description, and (optionally) an expected truth value.
- We updated several demo examples of CADP in order to illustrate the above extensions.

6.4.3. On-the-Fly Test Generation

Participants: Hubert Garavel, Radu Mateescu, Wendelin Serwe.

In the context of the collaboration with STMicroelectronics, we study techniques for testing if a (hardware) implementation is conform to a formal model described in LNT. Our approach is inspired by the theory of conformance testing [62], as implemented for instance in TGV [53] and JTorX [33]. We have developed two prototype tools to support this approach. The first tool implements a dedicated OPEN/CAESAR-compliant compiler for the particular asymmetric synchronous product between the model and the test purpose. The second tool, based on slightly extended generic components for graph manipulation (τ-compression, τ-confluence reduction, determinization) and resolution of Boolean equation systems, generates the complete test graph (CTG), which can be used to extract concrete test cases or to drive the test of the implementation. The principal advantage of our approach compared to existing tools is the use of LNT for describing test purposes, which facilitates the manipulation of data values.

In 2014, we developed a third prototype tool that takes as input a CTG and extracts either a single test case (randomly chosen or the first encountered one), or the set of *all* test cases. This prototype tool was used in the case study with STMicroelectronics (see § 6.5.1).

The test-generation tool TGV has been streamlined by removing some obsolete options and replacing a large part of its code by calls to the standard CADP libraries. TGV has been made faster, it now supports the latest version of the AUT format, and ensures that test purposes provided in the BCG format are deterministic. The manual page has been updated and completed.

6.4.4. Other Component Developments

Participants: Soraya Arias, Hubert Garavel, Frédéric Lang, Radu Mateescu.

The AUT textual format for CADP for storing LTSs was extended to support recent languages (such as LNT and the PseuCo language developed at Saarland University) that manipulate character-string values. The AUT format, which was defined in the late 80s, did not support such values. A new version 2014 of the AUT format has been defined, which solves this problem and maintains backward compatibility. All the CADP tools that read or write AUT files have been updated accordingly.

The BCG format of CADP for storing LTSs has been upgraded with the advent of a new version 1.2, which replaces version 1.1 released in 2009. New predefined types have been added to BCG to express the difference between unsigned and signed integers, and between character strings and untyped raw-data values. The new version of the BCG format is also more compact and now uses variable-length encoding for strings. The rules for label parsing of the BCG_WRITE interface have been extended, and BCG_IO now supports version 2014 of the AUT format. The intrinsic difficulty of these changes was to preserve the backward compatibility with the BCG files generated over the last twenty years.

To simplify the installation of CADP on Windows systems, we studied an alternative execution environment based on Gnuwin32 and MinGW/Msys rather than Cygwin. Preliminary changes have been brought to CADP scripts to undertake such a migration.

6.5. Real-Life Applications and Case Studies

6.5.1. ACE Cache Coherency Protocol

Participants: Abderahman Kriouile, Radu Mateescu, Wendelin Serwe.

In the context of a CIFRE convention with STMicroelectronics, we study system-level cache coherency, a major challenge faced in the current System-on-Chip architectures. Because of their increasing complexity (mainly due to the significant number of computing units), the validation effort using current simulation-based techniques grows exponentially. As an alternative, we study formal verification.

We focused on the ACE (AXI Coherency Extensions) cache coherency protocol, a system-level coherency protocol proposed by ARM [29]. In previous years, we developed a formal LNT model (about $3,400$ lines of LNT) of a system consisting of an ACE-based cache coherent interconnect, processors, and a main memory. The model is parametric and can be instantiated with different configurations (number of processors, number of cache lines, number of memory lines) and different sets of supported elementary ACE operations (currently, a representative subset of 15 operations), including an abstract operation that represents any other ACE operation. We handled the global requirements of the ACE specification using a constraint oriented programming style, i.e., by representing each global requirement as a dedicated process observing the global behavior and inhibiting incorrect executions. We also specified temporal properties expressing cache coherence, data integrity, and successful completion of each transaction. Note that the former property required to transform state-based properties into action-based ones, by adding information about the cache state to the actions executed by the cache.

In 2014, we exploited the formal model to improve the validation of the architecture under design at STMicroelectronics. In a first step, we studied the sanity (soundness and completeness) of an industrial interface verification unit, consisting of a list of so-called *formal checks*. After modeling each check in LNT, we used the BISIMULATOR tool to verify that each check is an overapproximation of the corresponding projection of the formal model. When we tried to establish that the parallel composition of all checks is an overapproximation of the projection of the formal model, we discovered a missing check (a particular channel did not occur in any of the checks).

In a second step, we studied the derivation of system level test cases, using a two-phase approach:

- In the first phase, abstract test cases were extracted automatically from the formal model using a prototype tool (see § 6.4). To circumvent the complexity of extracting test cases from the complete model, we proposed an iterative approach based on the automatic selection of a comprehensive set of interesting scenarios leading to LTSs of tractable size. The selection of the interesting scenarios relies on the counterexamples provided by the EVALUATOR model checker for the properties of coherence and data integrity.
- In the second phase, the abstract test cases were translated into the input format of an industrial test bench in charge of refining them into concrete test cases to be executed on the RTL (*Register Transfer Level*) description of the architecture under study. Experiments with manually translated abstract test cases led to the early discovery of bugs in commercial verification blocks, which could therefore be corrected before their use became critical in the development process.

The tests derived from the formal model increased the coverage of problematic features of some blocks used in the architecture. In particular, our approach was able to detect a limitation concerning data integrity 20 months before it was confirmed by classical methods, and our methodology provides all the scenarios triggering the limitation.

This work led to a publication accepted in an international conference [15]. Also, a large Petri net derived from our LNT model was provided as benchmark example for the Model Checking Contest.

6.5.2. Formal Verification of BPMN Processes

Participants: Radu Mateescu, Gwen Salaün, Lina Ye.

A business process is a set of structured, related activities that aims at fulfilling a specific organizational goal for a customer or market. An important metric when developing a business process is its degree of parallelism, i.e., the maximum number of tasks that are executable in parallel in that process. The degree of parallelism determines the peak demand on tasks, providing a valuable guide for the problem of resource allocation in business processes.

In 2014, we investigated how to automatically measure the degree of parallelism for business processes, described using the BPMN standard notation. To this aim, we defined a formal model for BPMN processes in terms of LTSs, which are obtained through an encoding in LNT. We then proposed an approach for automatically computing the degree of parallelism by using model checking of parameterized MCL formulas and dichotomic search. We developed a prototype tool for automating this check and we applied it successfully to more than one hundred BPMN processes.

This work led to a publication in an international conference [16].

6.5.3. Stability of Asynchronously Communicating Systems

Participants: Gwen Salaün, Lina Ye.

Analyzing communicating systems that interact asynchronously via reliable FIFO buffers is an undecidable problem. A typical approach is to check whether the system is bounded, and if not, the corresponding state space can be made finite by limiting the presence of communication cycles in behavioral models or by fixing buffer sizes.

We followed a different approach, which aims at analyzing communicating systems without restricting them by imposing any arbitrary bounds. These systems are likely to be unbounded and therefore result in infinite state spaces. We introduce a notion of stability and prove that once the system is stable for a specific buffer bound (called stability bound), it remains stable whatever larger bounds are chosen for the buffers. This enables us to check certain properties on the (finite-state) system obtained for the stability bound and to ensure that the system will preserve them whatever larger bounds are used for buffers.

We have also proven that computing the stability bound is in general undecidable, and we proposed a semi-algorithm that successfully computes the stability bounds for many typical examples of communicating systems using heuristics and equivalence checking. This work is described in a research report [27].

6.5.4. Deployment and Reconfiguration Protocols for Cloud Applications

Participants: Rim Abid, Gwen Salaün.

In the context of the OpenCloudware project (see § 8.1.1.1), we collaborate with Noël de Palma and Fabienne Boyer (University Joseph Fourier), Xavier Etchevers and Thierry Coupaye (Orange Labs) in the field of cloud computing applications, which are complex distributed applications composed of interconnected software components running on distinct virtual machines (VMs). Setting up, (re)configuring, and monitoring these applications involve intricate management protocols, which fully automate these tasks while preserving application consistency as well as some key architectural invariants.

In 2014, we extended the specification of the self-deployment protocol to support VM failures. This led to a publication in an international conference [11], of which an extended version is under preparation for submission to an international journal.

We also worked on the dynamic reconfiguration of cloud applications. As a first attempt, we proposed to design this protocol using a publish-subscribe communication model [32]. In 2014, we improved the protocol to also support VM failures, and drastically validated the corresponding LNT specification using model checking. A paper presenting these results was submitted to an international journal. In parallel, we studied a version of this protocol where the different participants interact asynchronously via FIFO buffers. This led to a publication in an international conference [10].

As a new line of work, we undertook the study of controller synthesis techniques for the coordination of autonomic managers in asynchronous environments. Our approach relies on an encoding into LNT and on the application of several operations on automata (synchronous products, hiding, reduction) for synthesizing the corresponding controller using CADP tools. We also proposed automated techniques for generating Java code from an abstract model of the controller. For validation purposes, we applied our approach to real-world three-tier Web applications and showed that the introduction of a controller allows one to avoid erroneous situations due to the absence of coordination between autonomic managers.

6.5.5. Networks of Programmable Logic Controllers

Participants: Hubert Garavel, Fatma Jebali, Jingyan Jourdan-Lu, Frédéric Lang, Eric Léo, Radu Mateescu.

In the context of the Bluesky project (see § 8.1.2.1), we study the software applications embedded on the PLCs (*Programmable Logic Controllers*) manufactured by Crouzet Automatismes. One of the objectives of Bluesky is to enable the rigorous design of complex control applications running on several PLCs connected by a network. Such applications are instances of GALS (*Globally Asynchronous, Locally Synchronous*) systems composed of several synchronous automata embedded on individual PLCs, which interact asynchronously by exchanging messages. A formal analysis of these systems can be naturally achieved by using the formal languages and verification techniques developed in the field of asynchronous concurrency.

For describing the applications embedded on individual PLCs, Crouzet provides a dataflow language with graphical syntax and synchronous semantics, equipped with an ergonomic user-interface that facilitates the learning and use of the language by non-experts. To equip the PLC language of Crouzet with functionalities for automated verification, the solution adopted in Bluesky was to translate it into GRL (see § 6.1.3), which enables the connection to testing and verification tools covering the synchronous and asynchronous aspects.

In 2014, we have developed a set of GRL libraries implementing about 40 of the function blocks present in the PLC programming tool of Crouzet, to facilitate the integration of GRL in the PLC software design process. These function blocks include (among others) logic and comparison functions, timers, triggers, and counters. These GRL libraries have been used to model large applications provided by Crouzet. The GRL2LNT and GRL.OPEN tools (see § 6.1.3) provide a direct connection to all verification functionalities of CADP, in particular model checking and equivalence checking.

Regarding model checking, we have studied existing work in the verification of synchronous systems and GALS systems. We have identified a set of typical patterns of temporal properties (e.g., deadlocks, safety, liveness) relevant for GALS systems. These property patterns have been specified using MCL and checked on a set of feature-rich GRL examples using GRL.OPEN and EVALUATOR.

Regarding equivalence checking, the purpose is to compare the behavior of a GALS system with its *service*, which represents its desired observable behavior, modulo a suitable equivalence relation. We have studied existing work in equivalence checking for GALS systems and we have investigated how to formally define the expected service of a GALS system at the appropriate level of expressiveness and abstraction, which requires a careful identification of the observable actions corresponding to the interactions between the GALS system and its physical environment. We have modeled several examples of GALS systems in GRL, and experimented the definition of appropriate services and their usage for equivalence checking by means of GRL.OPEN and BISIMULATOR.

The validation approach we promote, together with our colleagues from the LCIS laboratory (Valence) in the Bluesky project, led to a common publication in a national conference [21].

6.5.6. EnergyBus Standard for Connecting Electric Components

Participants: Hubert Garavel, Wendelin Serwe.

The EnergyBus [0] is an upcoming industrial standard for electric power transmission and management, based on the CANopen field bus. It is developed by a consortium assembling all major industrial players (such as Bosch, Panasonic, and Emtas) in the area of light electric vehicles (LEV); their intention is to ensure interoperability between all electric LEV components. At the core of this initiative is a universal plug integrating a CAN-Bus [0] with switchable power lines. The central and innovative role of the EnergyBus is to manage the safe electricity access and distribution inside an EnergyBus network.

In the framework of the European FP7 project SENSATION (see § 8.2.1.1) a formal specification in LNT of the main EnergyBus protocols is being developed by Alexander Graf-Brill and Holger Hermanns at Saarland University [48], with the active collaboration of CONVECS.

In 2014, our joint work with Saarland University on the modeling, verification, and test case generation for the EnergyBus standard led to a common publication [13].

6.5.7. Graphical User-Interfaces and Plasticity

Participants: Hubert Garavel, Frédéric Lang, Raquel Oliveira.

In the context of the Connexion project (see § 8.1.1.2) and in close collaboration with Gaëlle Calvary, Eric Ceret, and Sophie Dupuy-Chessa (IIHM team of the LIG laboratory), we study the formal description and validation of graphical user-interfaces using the most recent features of the CADP toolbox. The case study assigned to LIG in this project is a prototype graphical user-interface [36] designed to provide human operators with an overview of a running nuclear plant. The main goal of the system is to inform the operators about alarms resulting from faults, disturbances, or unexpected events in the plant. Contrary to conventional control rooms, which employ large desks and dedicated hardware panels for supervision, this new-generation interface uses standard computer hardware (i.e., smaller screen(s), keyboard, and mouse), thus raising challenging questions on how to best provide synthetic views of the plant status. Another challenge is to introduce plasticity in such interface, so as to enable several supervision operators, including mobile ones outside of the control room, to get accurate information in real time.

We formally specified this new-generation interface in LNT, encompassing not only the usual components traditionally found in graphical user-interfaces, but also a model of the physical world (namely, a nuclear reactor with various fault scenarios) and a cognitive model of a human operator in charge of supervising the plant. Also, several desirable properties of the interface have been expressed in MCL and verified on the LNT model using CADP. This led to a publication in an international conference [17].

In 2014, we continued our activity along several directions. The LNT specification was matured in various respects. As a result of several interactions with EDF, the specification was enhanced with a more realistic representation of the plant (currently $5,358$ lines of LNT code). Besides, new desirable properties of the user-interfaces emerged with the evolution of the formal model, making a total of seven complex properties formally specified in MCL.

We initiated an integration of our formal model with an industrial control room prototype, provided by a partner in the project. To this aim, several improvements were done in the formal specification, and the integration is currently in progress.

We started to address the introduction of plasticity in the formal specification, a challenge that was identified in 2013. Plasticity is the capacity of a user-interface to withstand variations of its context of use (i.e., platform, user, environment) while preserving usability. We proposed two approaches to introduce plasticity in the analysis. The first one introduces in the formal model a representation of a plasticity engine (responsible for user-interfaces adaptation) and applies model checking to verify its properties. The second approach consists in formally specifying several versions of the user-interfaces, derived from adaptation, and applying equivalence checking to verify similarity relations on the user-interface models.

[0] http://www.energybus.org
[0] http://www.can-cia.org

6.5.8. Fault-Tolerant Routing for Network-on-Chip Architectures

Participant: Wendelin Serwe.

Fault-tolerant architectures provide adaptivity for on-chip communications, but also increase the complexity of the design, so that formal verification techniques are needed to check their correctness. In collaboration with Chris Myers and Zhen Zhang (University of Utah, USA), we studied an extension of the link-fault tolerant Network-on-Chip (NoC) architecture introduced by Wu *et al* [67] that supports multiflit wormhole routing.

To keep the state space manageable, the formal LNT model of the routing algorithm was constructed in several steps, applying different abstractions (structural and related to data). This modeling process led to several insights. First, it led to the discovery of a package leakage path that could lead to the complete loss of a packet and a deadlock. This error in the design of an arbiter was corrected in the subsequent models. Second, a buffering capacity in an arbiter was found to be crucial; this insight also led to a redesign of the arbiters. The resultant changes on the router and arbiter models uncovered interesting symmetries. Finally, we studied how deadlock freedom and tolerance of a single-link fault can be verified for a NoC architecture.

This work led to a publication in an international conference [20].

6.5.8.1. Other Case Studies

The demo examples of CADP, which have been progressively accumulated since the origins of the toolbox, are a showcase for the multiple capabilities of CADP, as well as a test bed to assess the new features of the toolbox. In 2014, the effort to maintain and enhance these demos has been pursued. The progressive migration to LNT has continued, by translating certain demos from LOTOS to LNT. Many demos have been enriched with value-passing temporal formulas that illustrate the conciseness and expressiveness of MCL and the capabilities of the EVALUATOR 4 model checker. Finally, many demos have been shortened and made more readable by using the new features of SVL, especially the "`property`" and "`|=`" statements that allow formulas to be gathered in a single SVL file rather than disseminated in a collection of MCL or XTL files.

CORIDA Team

6. New Results

6.1. Highlights of the Year

The CORIDA team organized two scientific meetings in 2014.

The first workshop, “Observers for finite and infinite dimensional systems” in April 2014, gathered people working in the field of control theory for finite and infinite dimensional systems.

Ten speakers from France, India, Portugal and Germany were invited for the second workshop, “Workshop in Mathematical Fluid Dynamics", in November 2014.

6.2. Analysis and control of fluids and of fluid-structure interactions

In [42], we consider a two dimensional collision problem for a rigid solid immersed in a cavity filled with a perfect fluid. we investigate the asymptotic behavior of the Dirichlet energy associated to the solution of a Laplace Neumann problem as the distance between the solid and the cavity's bottom tends to zero. We prove that the solid always reaches the cavity in finite time. The contact occurs with non zero (real shock) or null velocity velocity (smooth landing), depending on the tangency exponent at the contact point. The proof is based on a suitable change of variables sending to infinity the cusp singularity at the contact. More precisely, the initial Laplace Neumann problem is transformed into a generalized Neumann problem set on a domain containing a horizontal strip, whose length goes to infinity as the the solid gets closer to the the cavity's bottom.

In [43], we investigate the geometric inverse problem of determining, from the knowledge of the DtN operator of the problem, the positions and the velocities of moving rigid solids in a bounded cavity filled with a perfect fluid. We assume that the solids are small disks moving slowly. Using an integral formulation, we first derive the asymptotic expansion of the DtN map as the diameters of the disks tend to zero. Then, combining a suitable choice of exponential type data and the DORT technique (which is usually used in inverse scattering for the detection of point-like scatterers), we propose a reconstruction method for the unknown positions and velocities.

In [22], Ana Leonor Silvestre (Lisbon, Portugal) and Takéo Takahashi analyze the system fluid-rigid body in the case of where the rigid body is a ball of “small radius”. More precisely, they consider the limit system as the radius goes to zero. They recover the Navier-Stokes system with a particle following the the velocity of the fluid.

In [14], Mehdi Badra (University of Pau) and Takéo Takahashi study the feedback stabilization of a system composed by an incompressible viscous fluid and a rigid body. They stabilize the position and the velocity of the rigid body and the velocity of the fluid around a stationary state by means of a Dirichlet control, localized on the exterior boundary of the fluid domain and with values in a finite dimensional space. The first result concerns weak solutions in the two-dimensional case, for initial data close to the stationary state. The method is based on general arguments for stabilization of nonlinear parabolic systems combined with a change of variables to handle the fact that the fluid domain of the stationary state and of the stabilized solution are different. This additional difficulty leads to the assumption that the initial position of the rigid body is the position associated to the stationary state. Without this hypothesis, they work with strong solutions, and to deal with compatibility conditions at the initial time, they use finite dimensional dynamical controls. They prove again that for initial data close to the stationary state, they can stabilize the position and the velocity of the rigid body and the velocity of the fluid.

In [15], Mehdi Badra (University of Pau) and Takéo Takahashi use the Fattorini criterion (more known as the Hautus criterion) to obtain the feedback stabilizability of general linear and nonlinear parabolic systems. They then consider flow systems described by coupled Navier-Stokes type equations (such as MHD system or micropolar fluid system) to obtain the stabilizability by only considering a unique continuation property of a stationary Stokes system.

In [36], we use geometric control theory to investigate the existence and the design of optimal strokes for swimmers in Stokes of potential flows.

6.3. Frequency domain methods for the analysis and control of systems governed by PDE's

In [20], we use microlocal analysis techniques to build artificial boundary conditions for reltivistic quantum dynamics.

In [11], we give a complete analysis of some new domain decomposition techniques and investigate their approximations for application in quantum physics.

In the chapter [28], we give an introduction to the modeling and the simulation of equilibrium states of Gross-Pitaevskii equations modeling Bose-Einstein condensates.

In [17], we give the basic methodology to use the software 3D GPELab for the simulation of Bose-Einstein condensates.

In [10], we develop a pseudo-spectral iterative method to compute equilibrium state of fast rotating Gross-Pitaevskii equations.

In [18], we develop a new approximation and implementation of a Magnetic-to-Electric operator for 3D-Maxwell equations.

In [13], we consider the inverse problem of determining the potential in the dynamical Schrödinger equation on the interval by the measurement on the boundary. Using the Boundary Control Method we first recover the spectrum of the problem from the observation at either left or right end points. Taking advantage of the one-dimensional configuration, we recover then the spectral function, reducing the problem to the classical one of determining the potential from the spectral function. This can be done by known methods. In order to handle more realistic situations, we also consider the case where only a finite number of eigenvalues are available and we prove the convergence of the reconstruction method as this number tends to infinity.

6.4. Observality, controllability and stabilization in the time domain

In [27], we dealt with the problem of the stabilization of a switched linear system, the feedback law being based on the optimization of a quadratic criterion. The Lyapunov function used for the design of this law defines a tight upper bound of the value of the cost for a quadratic optimization problem related to the system. Thus the obtained control law is sub-optimal.

In [19] we deal with the problem of the output stabilization of linear impulsive systems. These system are a mix of continuous and discrete-time system. An observer is synthesized and the stabilization is ensured through a feedback law which depends on the estimated state provided by the observer.

In [29], we consider the design an high gain observers for a class of continuous dynamical systems with discrete-time measurements. In this work, the measurement sampling time is considered to be variable. Moreover, the new idea of the proposed work is to synthesize an observer requiring the less knowledge as possible from the output measurements. This is done by using an updated sampling time observer. Under the global Lipschitz assumption, the asymptotic convergence of the observation error is established. As an application of this approach, an state estimation problem of an academic bioprocess is studied, and its simulation results are discussed.

In [26], we propose an MPC control scheme for a linear system with real-time constraints.

In [25] and [12], we use precise energy estimates to provide an upper bound on the error made when replacing the dynamics of an infinite dimensional conservative quantum system by a finite dimensional projection.

In [34], we give a set of sufficient conditions for approximate controllability of closed quantum systems when the dipolar approximation has to be replaced by a more realistic quadratic modeling.

In [35], we investigate the regularity of propagators of bilinear control systems and extend a celebrated negative result of Ball, Marsden and Slemrod.

In [16], we consider an infinite dimensional system modelling a boost converter connected to a load via a transmission line. The governing equations form a system coupling the telegraph partial differential equation with the ordinary differential equations modeling the converter. The coupling is given by the boundary conditions and the nonlinear controller we introduce. We design a nonlinear saturating control law using a Lyapunov function for the averaged model of the system. The main results give the well-posedness and stability properties of the obtained closed loop system.

CQFD Project-Team

6. New Results

6.1. Highlights of the Year

Creation of the Associate Team Inria: CDSS (2014-2016) with the University of Sao Paulo, Brasil.

6.2. Approximate Kalman–Bucy filter for continuous-time semi-Markov jump linear systems

Participants: Benoîte de Saporta, Eduardo Costa.

We propose a new numerical approximation of the Kalman–Bucy filter for semi-Markov jump linear systems. This approximation is based on the selection of typical trajectories of the driving semi-Markov chain of the process by using an optimal quantization technique. The main advantage of this approach is that it makes pre-computations possible. We derive a Lipschitz property for the solution of the Riccati equation and a general result on the convergence of perturbed solutions of semi-Markov switching Riccati equations when the perturbation comes from the driving semi-Markov chain. Based on these results, we prove the convergence of our approximation scheme in a general infinite countable state space framework and derive an error bound in terms of the quantization error and time discretization step. We employ the proposed filter in a magnetic levitation example with markovian failures and compare its performance with both the Kalman–Bucy filter and the Markovian linear minimum mean squares estimator. This work was presented at the international conference [37] and is submitted to an international journal [50].

6.3. Modeling and optimization of a launcher integration process

Participants: Benoîte de Saporta, François Dufour, Christophe Nivot.

We are interested in the optimization of a launcher integration process. It comprises several steps from the production of the subassemblies to the final launch. The four subassemblies go through various types of operations such as preparation, integration, control and storage. These operations are split up into three workshops. Due to possible breakdowns or staff issues, the time spent in each workshop is supposed random. So is the time needed to deliver the subassemblies, for similar reasons including e.g. shipping delays. We also have to deal with constraints related to the architecture of the assembly process itself. Indeed, we have to take into account waiting policies between workshops. The workshops may work in parallel but can be blocked if their output is not transferred to the next workshop in line. Storage capacity of output products is limited.

Our goal is finding the best rates of delivery of the subassemblies, the best choice of architecture (regarding stock capacities) and the best times when to stop and restart the workshops to be able to carry out twelve launches a year according to a predetermined schedule at minimal cost. To solve this problem, we choose a mathematical model particularly suitable for optimization with randomness: Markov decision processes (MDPs).

We have implemented a numerical simulator of the process based on the MDP model. It provides the fullest information possible on the process at any time. The simulator has first been validated with deterministic histories. Random histories have then been run with exponentially distributed delivery times for the subassemblies and several families of random laws for the time spent in each workshop. Using Monte Carlo simulations, we obtain the distribution of the launch times. Preliminary optimization results allow choosing stock capacities and delivery rates that satisfy the launch schedule. Work is still in progress concerning cost minimization. It was presented at Airbus internal PhD seminar in November 2014.

6.4. Numerical approximation for optimal stopping of MDP under partial observation

Participants: Benoîte de Saporta, François Dufour, Christophe Nivot.

We consider the optimal stopping problem for a continuous finite-dimensional state space Markov chain under partial observation. Our aim is to build a numerical approximation of the value function. To do so, we first translate the problem into the Partially Observed Markov Decision Process (POMDP) framework. Then, we define the equivalent fully observed Markov Decision Process (MDP) on an infinite dimensional state space. Finally, we proposed a discretization scheme based on the discretization of an underlying measure to obtain a finite dimensional problem and a discretization of the resulting state space to obtain a fully discrete model that is numerically tractable. We prove the convergence of the approximation procedure. This work is still in progress and was presented at the workshop [31]

6.5. Classification of EEG signals by evolutionary algorithm

Participants: Marie Chavent, Pierrick Legrand, Leonardo Trujillo.

The goal of this work is to predict the state of alertness of an individual by analyzing the brain activity through electroencephalographic data (EEG) captured with 58 electrodes. Alertness is characterized here as a binary variable that can be in a "normal" or "relaxed" state. We collected data from 44 subjects before and after a relaxation practice, giving a total of 88 records. After a pre-processing step and data validation, we analyzed each record and discriminate the alertness states using our proposed "slope criterion". Afterwards, several common methods for supervised classification (k nearest neighbors, decision trees (CART), random forests, PLS and discriminant sparse PLS) were applied as predictors for the state of alertness of each subject. The proposed "slope criterion" was further refined using a genetic algorithm to select the most important EEG electrodes in terms of classification accuracy. Results show that the proposed strategy derives accurate predictive models of alertness.

This work has been published in a book chapter [45].

6.6. Probabilistic low-rank matrix completion with adaptive spectral regularization algorithms

Participants: Marie Chavent, Adrien Todeschini.

We propose a novel class of algorithms for low rank matrix completion. Our approach builds on novel penalty functions on the singular values of the low rank matrix. By exploiting a mixture model representation of this penalty, we show that a suitably chosen set of latent variables enables to derive an EM algorithm to obtain a Maximum A Posteriori estimate of the completed low rank matrix. The resulting algorithm is an iterative soft-thresholded algorithm which iteratively adapts the shrinkage coefficients associated to the singular values.

This work is in collaboration with Francois Caron from University of Oxford. It has been presented in the national conference of the French Statistical Society of Statistics [41]

6.7. Variable selection to construct indicators of quality of life for data structured in groups

Participants: Marie Chavent, Amaury Labenne, Jérôme Saracco.

The analysis and measurement of quality of life may be made via two complementary approaches. The first one, based on survey of individuals, concerns the analysis of levels of life satisfaction. We focus here on the second one, based on national data, which analyses living conditions of people. The aim is to create composite indices of living conditions. According to authors, the components of quality of life are related to different themes (groups of variables): Family conditions", Employment", Housing",... For this purpose, dimension reduction methods are particularly suitable. Multiple Factor Analysis (MFA) is a method designed to handle data structured into groups of quantitative variables. In our study, each theme is composed of a group of quantitative and/or categorical variables. Since our data are naturally structured in groups of variables, we develop an extension of MFA for mixed data type, called MFAmix. Thus the principal components from MFAmix are our composite indices for measuring quality of life. However, the creation of these indices raises two questions. How many principal components keep to create indices? How select a limited number of variables to get similar indices for easier interpretation? We propose answers to these questions in this communication.

This work is in collaboration with Vanessa Kuentz from Irstea. It has been presented in the french meeting of the R users (Rencontres R) [40] and in the international conference COMPSTAT 2014 [36].

6.8. Efficiency of simulation in monotone hyper-stable queueing networks

Participants: Jonatha Anselmi, Bruno Gaujal.

We consider Jackson queueing networks with finite buffer constraints (JQN) and analyze the efficiency of sampling from their stationary distribution. In the context of exact sampling, the monotonicity structure of JQNs ensures that such efficiency is of the order of the 'coupling time' (or meeting time) of two extremal sample paths. In the context of approximate sampling, it is given by the 'mixing time'. Under a condition on the drift of the stochastic process underlying a JQN, which we call hyper-stability, in our main result we show that the coupling time is polynomial in both the number of queues and buffer sizes. Then, we use this result to show that the mixing time of JQNs behaves similarly up to a given precision threshold. Our proof relies on a recursive formula relating the coupling times of trajectories that start from network states having 'distance one', and it can be used to analyze the coupling and mixing times of other Markovian networks, provided that they are monotone. An illustrative example is shown in the context of JQNs with blocking mechanisms. This work has been published in an international journal; see [11].

6.9. Control of parallel non-observable queues: asymptotic equivalence and optimality of periodic policies

Participants: Jonatha Anselmi, Bruno Gaujal, Tommaso Nesti.

We consider a queueing system composed of a dispatcher that routes deterministically jobs to a set of non-observable queues working in parallel. In this setting, the fundamental problem is which policy should the dispatcher implement to minimize the stationary mean waiting time of the incoming jobs. We present a structural property that holds in the classic scaling of the system where the network demand (arrival rate of jobs) grows proportionally with the number of queues. Assume that each queue of type r is replicated k times and consider the set of policies that are periodic with period $k\sum_r p_r$ and such that exactly p_r jobs are sent in a period to each queue of type r. When $k \to \infty$, our main result shows that all the policies in this set are equivalent, in the sense that they yield the same mean stationary waiting time, and optimal, in the sense that no other policy having the same aggregate arrival rate to all queues of a given type can do better in minimizing the stationary mean waiting time. This property holds in a strong probabilistic sense. Furthermore, the limiting mean waiting time achieved by our policies is a convex function of the arrival rate in each queue, which facilitates the development of a further optimization aimed at solving the fundamental problem above for large systems. This work has been accepted for publication in the international journal “Stochastic Systems”, the flagship journal of the INFORMS Applied Probability Society; see [46].

6.10. The economics of the cloud: price competition and congestion

Participants: Jonatha Anselmi, Danilo Ardagna, Jonh C.s. Lui, Adam Wierman, Yunjian Xu, Zichao Yang.

This work proposes a model to study the interaction of price competition and congestion in the cloud computing marketplace. Specifically, we propose a three-tier market model that captures a marketplace with users purchasing services from Software-as-Service (SaaS) providers, which in turn purchase computing resources from either Provider-as-a-Service (PaaS) providers or Infrastructure-as-a-Service (IaaS) providers. Within each level, we define and characterize competitive equilibria. Further, we use these characterizations to understand the relative profitability of SaaSs and PaaSs/IaaSs, and to understand the impact of price competition on the user experienced performance, i.e., the 'price of anarchy' of the cloud marketplace. Our results highlight that both of these depend fundamentally on the degree to which congestion results from shared or dedicated resources in the cloud. This work has been submitted to an international journal. A preliminary has been published in [10].

6.11. Generalized Nash Equilibria for Platform-as-a-Service Clouds

Participants: Jonatha Anselmi, Danilo Ardagna, Mauro Passacantando.

Cloud computing is an emerging technology that allows to access computing resources on a pay-per-use basis. The main challenges in this area are the cient performance management and the energy costs minimization. In this work we model the service provisioning problem of Cloud Platform-as-a-Service systems as a Generalized Nash Equilibrium Problem and show that a potential function for the game exists. Moreover, we prove that the social optimum problem is convex and we derive some properties of social optima from the corresponding Karush-Kuhn-Tucker system. Next, we propose a distributed solution algorithm based on the best response dynamics and we prove its convergence to generalized Nash equilibria. Finally, we numerically evaluate equilibria in terms of their efficiency with respect to the social optimum of the Cloud by varying our algorithm initial solution. Numerical results show that our algorithm is scalable and very efficient and thus can be adopted for the run-time management of very large scale systems. This work has been published in an international journal; see [12].

6.12. Stochastic approximations of constrained discounted Markov decision processes

Participants: Francois Dufour, Tomas Prieto-Rumeau.

We consider a discrete-time constrained Markov decision process under the discounted cost optimality criterion. The state and action spaces are assumed to be Borel spaces, while the cost and constraint functions might be unbounded. We are interested in approximating numerically the optimal discounted constrained cost. To this end, we suppose that the transition kernel of the Markov decision process is absolutely continuous with respect to some probability measure μ. Then, by solving the linear programming formulation of a constrained control problem related to the empirical probability measure μ_n of μ, we obtain the corresponding approximation of the optimal constrained cost. We derive a concentration inequality which gives bounds on the probability that the estimation error is larger than some given constant. This bound is shown to decrease exponentially in n. Our theoretical results are illustrated with a numerical application based on a stochastic version of the Beverton-Holt population model. This work has been published in Journal of Mathematical Analysis and applications: [27].

6.13. Non-Parametric Estimation of the Conditional Distribution of the Interjumping Times for Piecewise-Deterministic Markov Processes

Participants: Romain Azais, Francois Dufour, Anne Gegout-Petit.

We study a nonparametric method for estimating the conditional density associated to the jump rate of a piecewise-deterministic Markov process. In our framework, the estimation needs only one observation of the process within a long time interval. Our method relies on a generalization of Aalen?s multiplicative intensity model. We prove the uniform consistency of our estimator, under some reasonable assumptions related to the primitive characteristics of the process. A simulation study illustrates the behavior of our estimator. This work has been published in Scandinavian Journal of Statistics: [15].

6.14. Approximation of average cost Markov decision processes using empirical distributions and concentration inequalities

Participants: Francois Dufour, Tomas Prieto-Rumeau.

We consider a discrete-time Markov decision process with Borel state and action spaces, and possibly unbounded cost function. We assume that the Markov transition kernel is absolutely continuous with respect to some probability measure μ. By replacing this probability measure with its empirical distribution μ_n for a sample of size n, we obtain a finite state space control problem, which is used to provide an approximation of the optimal value and an optimal policy of the original control model. We impose Lipschitz continuity properties on the control model and its associated density functions. We measure the accuracy of the approximation of the optimal value and an optimal policy by means of a non-asymptotic concentration inequality based on the 1–Wasserstein distance between μ and μ_n. Obtaining numerically the solution of the approximating control model is discussed and an application to an inventory management problem is presented. This work has been published in Stochastics An International Journal of Probability and Stochastic Processes: [26].

6.15. Piecewise Deterministic Markov Processes based approach applied to an offshore oil production system

Participants: Huilong Zhang, Fares Innal, François Dufour, Yves Dutuit.

This work is keeping with the topic of two papers which treated dynamic reliability problems and were presented in previous conferences. Its aim is to confirm the potentialities of a method which combines the high modeling ability of the piecewise deterministic processes and the great computing power inherent to the Monte Carlo simulation. This method is now applied to a simplified but realistic offshore oil production system which is a hybrid system combining continuous-time and discrete-time dynamics. The results thus obtained have been compared with those given by an ad hoc Petri net model for comparison and validation purposes. This work has been published in an international journal; see [29].

6.16. Optimal Trajectories for Underwater Vehicles by Quantization and Stochastic control

Participants: Huilong Zhang, Benoîte de Saporta, François Dufour, Dann Laneuville, Adrien Nègre.

We propose in this paper a numerical method which computes the trajectory of a vehicle subject to some mission objectives. The method is applied to a submarine whose goal is to best detect one or several targets (we consider signal attenuation due to acoustic propagation) or/and to minimize its own detection range perceived by the other targets. Our approach is based on dynamic programming of a finite horizon Markov decision process. The position and the velocity of the targets are supposed to be known only up to a random estimation error, as a Kalman type filter is used to estimate these quantities from the measurements given by the on board sonar. We also take into account the information on the environment through a sound propagation code. A quantization method is applied to fully discretize the problem and solve it numerically. This work is still in progress and was presented at the international conference [39].

6.17. Multi-Objective Design and Maintenance Optimization of the Heated Hold-Up Tank Modeled by Piecewise Deterministic Markov Processes

Participants: Huilong Zhang, Yanfu Li.

We propose a numerical method for the optimal design and maintenance for the heated hold-up tank system. A multi-objective problem is framed to consider simultaneously the objectives of maximizing the operation profit and maximizing the reliability. The system consists of a tank containing a fluid whose level is controlled by three components: two inlet pumps and one outlet valve. A thermal power source heats up the fluid. The failure rates of the components depend on the temperature, the position of the three components monitors the liquid level in the tank and the liquid level determines the temperature. We model the system by a piecewise deterministic Markov process. To find the solution of the optimal maintenance interval, the non-dominated sorting genetic algorithm-II (NSGA-II) is used. This work is still in progress and was presented at the international conference [42].

6.18. Conditional quantile estimation through optimal quantization

Participants: Isabelle Charlier, Jérôme Saracco.

We use quantization to construct a nonparametric estimator of conditional quantiles of a scalar response Y given a d-dimensional vector of covariates X. First we focus on the population level and show how optimal quantization of X, which consists in discretizing X by projecting it on an appropriate grid of N points, allows to approximate conditional quantiles of Y given X. We show that this approximation is arbitrarily good as N goes to infinity and provide a rate of convergence for the approximation error. Then we turn to the sample case and define an estimator of conditional quantiles based on quantization ideas. We prove that this estimator is consistent for its fixed-N population counterpart. The results are illustrated on a numerical example. This work is in collaboration with Davy Paindaveine from Université Libre de Bruxelles. It has been presented in the national conference of the French Statistical Society of Statistics [35] and in the international conference on computational statistics [34].

6.19. Conditional quantile estimator based on optimal quantization: from theory to practice

Participants: Isabelle Charlier, Jérôme Saracco.

[21] recently introduced a promising nonparametric estimator of conditional quantiles based on optimal quantization, but almost exclusively focused on its theoretical properties. We now discuss its practical implementation (by proposing in particular a method to properly select the corresponding smoothing parameter, namely the number of quantizers) and (ii) we investigate how its finite-sample performances compare with those or classical kernel of nearest-neighbor competitors. Monte Carlo studies show that the quantization-based estimator competes well in all cases (in terms of mean squared errors) and tends to dominate its competitors as soon as the covariate is not uniformly distributed over its support. We also apply our approach to a real data set. While most of the paper focuses on the case of a univariate covariate, we also briefly discuss the multivariate case and provide an illustration for bivariate regressors. This work is in collaboration with Davy Paindaveine from Université Libre de Bruxelles. It has been presented in the national conference of the French Statistical Society of Statistics [35] and in the international conference on computational statistics [34].

6.20. QuantifQuantile : an R package for performing quantile regression trough optimal quantization

Participants: Isabelle Charlier, Jérôme Saracco.

Quantile regression allows to assess the impact of some covariate X on a response Y . An important application is the construction of reference curves and conditional prediction intervals for Y . Recently, [21] developed a new nonparametric quantile regression method based on the concept of optimal quantization. We now describe an R package, called QuantifQuantile, that allows to perform quantization-based quantile regression. We describe the various functions of the package and provide examples. This work is in collaboration with Davy Paindaveine from Université Libre de Bruxelles. It has been presented in the national conference on the R software [43].

6.21. Transcriptome profile analysis reveals specific signatures of pollutants in Atlantic eels

Participant: Jérôme Saracco.

Identifying specific effects of contaminants in a multi-stress field context remain a challenge in ecotoxicology. In this context, "omics" technologies, by allowing the simultaneous measurement of numerous biological endpoints, could help unravel the in situ toxicity of contaminants. In this study, wild Atlantic eels were sampled in 8 sites presenting a broad contamination gradient in France and Canada. The global hepatic transcriptome of animals was determined by RNA-Seq. In parallel, the contamination level of fish to 8 metals and 25 organic pollutants was determined. Factor analysis for multiple testing was used to identify genes that are most likely to be related to a single factor. Among the variables analyzed, arsenic (As), cadmium (Cd), lindane (γ-HCH) and the hepato-somatic index (HSI) were found to be the main factors affecting eel's transcriptome. Genes associated with As exposure were involved in the mechanisms that have been described during As vasculotoxicity in mammals. Genes correlated with Cd were involved in cell cycle and energy metabolism. For γ-HCH, genes were involved in lipolysis and cell growth. Genes associated with HSI were involved in protein, lipid and iron metabolisms. Our study proposes specific gene signatures of pollutants and their impacts in fish exposed to multi-stress conditions.

This work is in collaboration with G. Durrieu from Vannes University and R. Coudret. It will be published in Ecotoxicology [17].

6.22. Comparaison of kernel density estimators with assumption on number of modes : application on environmental monitoring data

Participant: Jérôme Saracco.

A data-driven bandwidth choice for a kernel density estimator called critical bandwidth is investigated. This procedure allows the estimation to have as many modes as assumed for the density to estimate. Both Gaussian and uniform kernels are considered. For the Gaussian kernel, asymptotic results are given. For the uniform kernel, an argument against these properties is mentioned. These theoretical results are illustrated with a simulation study that compares the kernel estimators that rely on critical bandwidth with another one that uses a plug-in method to select its bandwidth. An estimator that consists in estimates of density contour clusters and takes assumptions on number of modes into account is also considered. Finally, the methodology is illustrated using environment monitoring data.

This work is in collaboration with G. Durrieu from Vannes University and R. Coudret. It will be published in Communication in Statistics - Simulation and Computation [28].

6.23. A new sliced inverse regression method for multivariate response

Participant: Jérôme Saracco.

A semiparametric regression model of a q-dimensional multivariate response y on a p-dimensional covariate x is considered. A new approach is proposed based on sliced inverse regression (SIR) for estimating the effective dimension reduction (EDR) space without requiring a prespecified parametric model. The convergence at rate square root of n of the estimated EDR space is shown. The choice of the dimension of the EDR space is discussed. Moreover, a way to cluster components of y related to the same EDR space is provided. Thus, the proposed multivariate SIR method can be used properly on each cluster instead of blindly applying it on all components of y. The numerical performances of multivariate SIR are illustrated on a simulation study. Applications to a remote sensing dataset and to the Minneapolis elementary schools data are also provided. Although the proposed methodology relies on SIR, it opens the door for new regression approaches with a multivariate response.

This work is in collaboration with S. Girard from Inria MISTIS team and R. Coudret. It is published in CSDA [23].

6.24. An introduction to dimension reduction in nonparametric kernel regression

Participant: Jérôme Saracco.

Nonparametric regression is a powerful tool to estimate nonlinear relations between some predictors and a response variable. However, when the number of predictors is high, nonparametric estimators may suffer from the curse of dimensionality. In this chapter, we show how a dimension reduction method (namely Sliced Inverse Regression) can be combined with nonparametric kernel regression to overcome this drawback. The methods are illustrated both on simulated datasets as well as on an astronomy dataset using the R software [51].

This work is in collaboration with S. Girard from Inria MISTIS team .

6.25. Hidden Markov Model for the detection of a degraded operating mode of optronic equipment

Participant: Jérôme Saracco.

As part of optimizing the reliability, Thales Optronics now includes systems that examine the state of its equipment. The aim of this paper is to use hidden Markov Model to detect as soon as possible a change of state of optronic equipment in order to propose maintenance before failure. For this, we carefully observe the dynamic of a variable called "cool down time" and noted Tmf, which reflects the state of the cooling system. Indeed, the Tmf is an indirect observation of the hidden state of the system. This one is modelled by a Markov chain and the Tmf is a noisy function of it. Thanks to filtering equations, we obtain results on the probability that an appliance is in degraded state at time t, knowing the history of the Tmf until this moment. We have evaluated the numerical behavior of our approach on simulated data. Then we have applied this methodology on our real data and we have checked that the results are consistent with the reality. This method can be implemented in a HUMS (Health and Usage Monitoring System). This simple example of HUMS would allow the Thales Optronics Company to improve its maintenance system. This company will be able to recall appliances which are estimated to be in degraded state and do not control to soon those estimated in stable state.

This work is in collaboration with A. Gegout-Petit from Lorraine University. It is published in Journal de la SFdS [19].

6.26. On the asymptotic behavior of the Nadaraya-Watson estimator associated with the recursive SIR method

Participant: Jérôme Saracco.

We investigate the asymptotic behavior of the Nadaraya-Watson estimator for the estimation of the regression function in a semiparametric regression model. On the one hand, we make use of the recursive version of the sliced inverse regression method for the estimation of the unknown parameter of the model. On the other hand, we implement a recursive Nadaraya-Watson procedure for the estimation of the regression function which takes into account the previous estimation of the parameter of the semiparametric regression model. We establish the almost sure convergence as well as the asymptotic normality for our Nadaraya-Watson estimator. We also illustrate our semiparametric estimation procedure on simulated data.

This work is in collaboration with B. Bercu from Bordeaux University and T.M.N Nguyen. It is published in Statistics [20].

6.27. Evolving Genetic Programming Classifiers with Novelty Search

Participants: Enrique Naredo, Leonardo Trujillo, Pierrick Legrand.

Novelty Search (NS) is a unique approach towards search and optimization, where an explicit objective function is replaced by a measure of solution novelty to provide the selective pressure in an artificial evolutionary system. However, NS has been mostly used in evolutionary robotics, while it's applicability to classic machine learning problems has been mostly unexplored. This work presents a NS-based Genetic Programming (GP) algorithm for supervised classification, with the following noteworthy contributions. It is shown that NS can solve real-world classification tasks, validated over several commonly used benchmarks. These results are made possible by using a domain-specific behavioral descriptor, closely related to the concept of semantics in GP. Moreover, two new variants of the NS algorithm are proposed, Probabilistic NS (PNS) and a variant of Minimum Criterion NS (MCNS). The former models the behavior of each solution as a random vector, eliminating all the NS parameters and reducing the computational overhead of the traditional NS algorithm; the latter uses a standard objective function to constrain the search and bias the process towards high performance solutions. The paper also discusses the effects of NS on an important GP phenomenon, bloat. In particular, results indicate that some variants of the NS approach can have a beneficial effect on the search process by curtailing code growth. See [52].

6.28. Detecting mental states of alertness with genetic algorithm variable selection

Participants: Laurent Vezard, Pierrick Legrand, Marie Chavent, Frédérique Faïta, Léonardo Trujillo.

The objective of the present work is to develop a method that is able to automatically determine mental states of vigilance; i.e., a person's state of alertness. Such a task is relevant to diverse domains, where a person is expected or required to be in a particular state of mind. For instance, pilots and medical staff are expected to be in a highly alert state and the proposed method could help to detect possible deviations from this expected state. This work poses a binary classification problem where the goal is to distinguish between a “relaxed" state and a baseline state (“normal") from the study of electroencephalographic signals (EEG) collected with a small number of electrodes. The EEG of 58 subjects in the two alertness states (116 records) were collected via a cap with 58 electrodes. After a data validation step, 19 subjects were retained for further analysis. A genetic algorithm was used to select a subset of electrodes. Common spatial pattern (CSP) coupled to linear discriminant analysis (LDA) was used to build a decision rule and thus predict the alertness of the subjects. Different subset sizes were investigated and the best compromise between the number of selected electrodes and the quality of the solution was obtained by considering 9 electrodes. Even if the present approach is costly in computation time (GA search), it allows to construct a decision rule that provides an accurate and fast prediction of the alertness state of an unseen individual. See [45], [54].

6.29. A comparison of fitness-case sampling methods for Symbolic Regression

Participants: Yuliana Martinez, Léonardo Trujillo, Enrique Naredo, Pierrick Legrand.

The canonical approach towards fitness evaluation in Genetic Programming (GP), is to use a static training set to determine fitness, based on a cost function (root-mean-squared error) averaged over all cases. However, motivated by different goals, researchers have recently proposed several techniques that focus selective pressure on a subset of fitnesscases at each generation. These approaches can be described as fitnesscase sampling techniques, where the training set is sampled, in someway, to determine fitness. This paper shows a comprehensive evaluation of some sampling methods using benchmark problems and real-world problems. The algorithms considered here are Interleaved Sampling, Random Interleaved Sampling, Lexicase Selection and a new sampling technique is proposed called Keep-Worst Interleaved Sampling (KW-IS). The algorithms are extensively evaluated based on test performance, overfitting and bloat. Results suggest that sampling techniques can improve performance based on testing error, bloat and overfitting compared to standard GP. Some of the best results were achieved by Lexicase Selection and Keep Worse-Interleaved Sampling which obtained good results in overfitting and bloat effect. Results also show that on these problems overfitting correlates strongly with bloating and exhibits a good compromise among the considered performance measures.

6.30. Geometric Semantic Genetic Programming with Local Search

Participants: Emigdio Z. Flores, Léonardo Trujillo, Leonardo Vanneshi, Sara Silva, Pierrick Legrand.

Since its introduction, Geometric Semantic Genetic Programming (GSGP) has aroused the interest of numerous researchers and several studies have demonstrated that GSGP is able to effectively optimize training data by means of small variation steps, that also have the effect of limiting overfitting. In order to speed up the search process, in this paper we propose a system that integrates a local search strategy into GSGP (called GSGP-LS). Furthermore, we present a hybrid approach, that combines GSGP and GSGP-LS, aimed at exploiting both the optimization speed of GSGP-LS and the ability to limit overfitting of GSGP. The experimental results we present, performed on a set of complex real-life applications, show that GSGP-LS achieves the best training fitness while converging very quickly, but severely overfits; GSGP converges very slowly, but is basically not affected by overfitting. The best overall results were achieved with the hybrid approach, allowing the search to converge quickly, while also exhibiting a noteworthy ability to limit overfitting. These results are encouraging, and suggest that future GSGP algorithms should focus on finding the correct balance between the greedy optimization of a local search strategy and the more robust geometric semantic operators.

CRYPT Team

4. New Results

4.1. Highlights of the Year

The team published [20] improved single-key attacks on reduced-round AES: AES is currently the most widespread block cipher standard, it is implemented in Intel processors.

The team also showed [18] how to speed-up a well-known public-key cryptanalysis technique: finding small roots of univariate polynomial congruences. This technique is used to break special cases of the RSA cryptosystem.

Phong Nguyen was Program co-Chair of the 33rd IACR Eurocrypt Conference (EUROCRYPT 2014) [22].

CTRL-A Exploratory Action

6. New Results

6.1. Highlights of the Year

We have been invited to participate to the organization of events, which highlight our active presence in the scientific life in the two domains which we are bridging :

- autonomic computing: Eric Rutten is PC member, as well as workshops chair, of the 12th IEEE International Conference on Autonomic Computing, ICAC 2015 (http://icac2015.imag.fr/), and PC co-chair of the 3rd IEEE International Conference on Cloud and Autonomic Computing, CAC 2015 (http://autonomic-conference.org/), the two major conferences on the topic.
- control: Eric Rutten is organizer of a special session on discrete control for computing at the 12th IFAC - IEEE International Workshop on Discrete Event Systems, WODES 2014 (http://wodes2014.lurpa.ens-cachan.fr/), the main conference specialized in Discrete Event Systems, ; he is on the IFAC Technical Committee 1.3 on Discrete Event and Hybrid Systems, (http://tc.ifac-control.org/1/3/) and on the IEEE Control Systems Society Discrete Event Systems Technical Committee (http://discrete-event-systems.ieeecss.org).

6.2. Discrete control and reactive language support

Participants: Gwenaël Delaval, Eric Rutten, Stéphane Mocanu.

Concerning language support, we have designed and implemented BZR, a mixed imperative/declarative programming language: declarative contracts are enforced upon imperatively described behaviors (see 5.1). The semantics of the language uses the notion of Discrete Controller Synthesis (DCS) [5]. We target the application domain of adaptive and reconfigurable systems: our language can serve programming closed-loop adaptation controllers, enabling flexible execution of functionalities w.r.t. changing resource and environment conditions. DCS is integrated into a programming language compiler, which facilitates its use by users and programmers, performing executable code generation. The tool is concretely built upon the basis of a reactive programming language compiler, where the nodes describe behaviors that can be modeled in terms of transition systems. Our compiler integrates this with a DCS tool [3]. This work is done in close cooperation with the Inria team Sumo at Inria Rennes (H. Marchand). Ongoing work concerns aspects of compilation and debugging and logico-numeric extension of BZR based on the ReaX tool developed at Inria Rennes in the framework of the ANR Ctrl-Green project (see 8.2.1).

We are also currently working on on combining maximally permissive discrete control with runtime mechanisms for choosing between valid control values, involving e.g. a classical controller or stochastic aspects ; and on exploring the notion of adaptive discrete control, which is yet an open question in discrete control in contrast to the well-known adaptive continuous control.

Another activity related to discrete control is or work with Leiden University and CWI (N. Khakpour, now at Linnaeus U., and F. Arbab) on enforcing correctness of the behavior of an adaptive software system during dynamic adaptation is an important challenge along the way to realize correct adaptive systems. In this research, we model adaptation as a supervisory control problem and synthesize a controller that guides the behavior of a software system during adaptation. The system during adaptation is modeled using a graph transition system and properties to be enforced are specified using an automaton. To ensure correctness, we then synthesize a controller that imposes constraints on the system during adaptation [14].

6.3. Design and programming

6.3.1. Component-based approaches

Participants: Frederico Alvares, Eric Rutten.

Component-based architectures have shown to be very suited for self-adaptation purposes, not only because of their intrinsic characteristics like reusability and modularity, but also as virtue of their dynamical reconfiguration capabilities. The issue, nevertheless, remains that adaptation behaviors are generally conceived by means of fine-grained reconfiguration actions from the very initial configurations. This way, besides the complexity in managing large-sized architectures, the space of reachable configurations is not know in advance, which prevents ensuring well-mastered adaptive behaviours. We address this problem by designing Ctrl-F, a domain-specific language which objective is to provide high-level support for describing adaptation behaviors and policies in component-based architectures. The proposed language lies on synchronous reactive programming, which means that it benefits of an entire environment and formal tooling allowing for the verification and control of reconfigurations. We show the applicability of Ctrl-F by first integrating it to FraSCAti, a Service Component Architecture middleware platform, and then by applying it to Znn.com, a well known self-adaptive case study.

We work on the topic in cooperation with the Spirals Inria team at Inria Lille (L. Seinturier). It constitutes a follow-up on previous work in the ANR Minalogic project MIND, industrializing the Fractal component-based framework, with a continuation of contacts with ST Microelectronics (V. Bertin). Our integration of BZR and Fractal [4], [2] is at the basis of our current work. On a related topic, we are also starting a cooperation on introducing reactive control in hierarchical autonomic architectures, with A. Diaconescu and E. Najm at TelecomParisTech.

6.3.2. Rule-based systems

Participants: Julio Cano, Adja Sylla, Gwenaël Delaval, Eric Rutten.

Event-Condition-Action (ECA) rules are a widely used language for the high level specification of controllers in adaptive systems, such as Cyber-Physical Systems and smart environments, where devices equipped with sensors and actuators are controlled according to a set of rules. The evaluation and execution of every ECA rule is considered to be independent from the others, but interactions of rule actions can cause the system behaviors to be unpredictable or unsafe. Typical problems are in redundancy of rules, inconsistencies, circularity, or application- dependent safety issues. Hence, there is a need for coordination of ECA rule-based systems in order to ensure safety objectives. We propose a tool-supported method for verifying and controlling the correct interactions of rules, relying on formal models related to reactive systems, and Discrete Controller Synthesis (DCS) to generate correct rule controllers [12].

We work on this topic in cooperation with CEA LETI/DACLE (L. Gurgen) and target the application and experimentation domain of smart environment in the Internet of Things [11].

Another complementary direction on which we are starting a cooperation with CEA LETI/DACLE is the topic of a high-level language for safe rule-based programming in the LINC platform: the PhD of Adja Sylla on this topic will be co-advised with F. Pacull and M. Louvel at CEA.

6.4. Infrastructure-level support

6.4.1. Autonomic Cloud and Big-Data systems

This activity continues work started several years ago in the Sardes Inria-team, before it split into Erods (at LIG) and Ctrl-A (at Inria).

6.4.1.1. Coordination in multiple-loop autonomic Cloud systems

Participants: Soguy Gueye, Gwenaël Delaval, Stéphane Mocanu, Bogdan Robu, Eric Rutten.

Complex computing systems are increasingly self-adaptive, with an autonomic computing approach for their administration. Real systems require the co-existence of multiple autonomic management loops, each complex to design. However their uncoordinated co-existence leads to performance degradation and possibly to inconsistency. There is a need for methodological supports facilitating the coordination of multiple autonomic managers. We address this problem in the context of the ANR project Ctrl-Green (see 8.2.1), in cooperation with LIG (N. de Palma) in the framework of the PhD of S. Gueye. We propose a method focusing on the discrete control of the interactions of managers [7] [9]. We follow a component-based approach and explore modular discrete control, allowing to break down the combinatorial complexity inherent to the state-space exploration technique [13]. This improves scalability of the approach and allows constructing a hierarchical control. It also allows re-using complex managers in different contexts without modifying their control specifications. We build a component-based coordination of managers, with introspection, adaptivity and reconfiguration. We validate our method on a multiple-loop multi-tier system.

We are currently working on the distributed execution of modular controllers and on considering more control objectives, beyond purely discrete or logical ones, evaluating the new tool ReaX developed at Inria Rennes (Sumo) (see 6.2) and exploring continuous or stochastic control of servers provisioning.

6.4.1.2. Control for Big data

Participants: Bogdan Robu, Mihaly Berekmeri, Nicolas Marchand.

To deal with the issue of ensuring performance constraints while also minimizing costs in systems for Big Data analytics based on the parallel programming paradigm MapReduce, we propose a control theoretical approach, based on techniques that have already proved their usefulness for the control community. We develop an algorithm to create the first linear dynamic model for a Big Data MapReduce system, running a concurrent workload. Furthermore we identify two major performance constraint use cases: relaxed-minimal resource and strict performance constraints. For the first case we developed a feedback control mechanism and, to minimize the number of control actuations, an event-based feedback controller. For the second case we add a feedforward controller that efficiently suppresses the effects of large workload size variations. The work is validated in a simulated Matlab environment build at GIPSA-lab and online on a real 60 node MapReduce cluster (part of GRID 500), running a data intensive Business Intelligence workload. Our experiments demonstrate the success of the control strategies employed in assuring service time constraints [17], [18].

This work is performed in cooperation with LIG (S. Bouchenak) in the framework of the PhD of M. Berekmeri.

6.4.2. Reconfiguration control in DPR FPGA

Participant: Eric Rutten.

Dynamically reconfigurable hardware has been identified as a promising solution for the design of energy efficient embedded systems. However, its adoption is limited by the costly design effort including verification and validation, which is even more complex than for non dynamically reconfigurable systems. We work on this topic in the context of a ensign environment, developed in the framework of the ANR project Famous, in cooperation with LabSticc in Lorient and Inria Lille (DaRT team) [10]. We propose a tool-supported formal method to automatically design a correct-by-construction control of the reconfiguration. By representing system behaviors with automata, we exploit automated algorithms to synthesize controllers that safely enforce reconfiguration strategies formulated as properties to be satisfied by control. We design generic modeling patterns for a class of reconfigurable architectures, taking into account both hardware architecture and applications, as well as relevant control objectives. We validate our approach on two case studies implemented on FPGAs [1].

We are currently valorizing results in more publications, and extending the use of control techniques by evaluating the new tool ReaX developed at Inria Rennes (Sumo) in the framework of the ANR Ctrl-Green project (see 6.2 and 8.2.1).

6.4.3. Autonomic memory management in HPC

Participants: Naweiluo Zhou, Gwenaël Delaval, Bogdan Robu, Eric Rutten.

Concurrent programs need to manage the time trade-off between synchronization and computing. A high concurrency level may decrease computing time but at the same time increase synchronization cost among threads. The traditional way to handle synchronization problems is through implementing locks. However locks suffer from the likelihood of deadlocks, vulnerability to failures, faults etc.. Software Transactional Memory (STM) has emerged as a promising technique to address synchronization issues through transactions. In STM, blocks of instructions accessing the shared data are wrapped into transactions. In STM each transaction executes speculatively, and conflicts may be aroused when two transactions are trying to modify the same area simultaneously. A way to reduce conflicts is by adjusting concurrency levels. A suitable concurrency level can maximize program performance. However, there is no universal rule to decide the best concurrency level for a program from an offline view. Hence, it becomes necessary to adopt a dynamical tuning strategy to better manage a STM system, so that a program can achieve a better performance. In the context of the action-team HPES of the Labex Persyval-lab[0] (see 8.1), we explore the autonomic computing approach and control techniques to address these runtime tuning problems as a feedback control loop to automate the choices of concurrency levels, conflict management policies, and other parameters, with the objective of optimizing program execution time. This work is performed in cooperation with LIG (J.F. Méhaut) in the framework of the PhD of N. Zhou.

6.4.4. Control of smart environments

Participants: Julio Angel Cano Romero, Mengxuan Zhao, Eric Rutten, Hassane Alla [Gipsa-lab].

6.4.4.1. Generic supervision architecture

New application domains of control, such as in the Internet of Things (IoT) and Smart Environments, require generic control rules enabling the systematization and the automation of the controller synthesis. We are working on an approach for the generation of Discrete Supervisory Controllers for these applications. A general modeling framework is proposed for the application domain of smart home. We formalize the design of the environment manager as a Discrete Controller Synthesis (DCS) problem, w.r.t. multiple constraints and objectives, for example logical issues of mutual exclusion, bounding of power peaks. We validate our models and manager computations with the BZR language and an experimental simulator [15]. This work is performed in cooperation with Orange labs (G. Privat) in the framework of the Cifre PhD of M. Zhao.

6.4.4.2. Rule-based specification

In the Internet of things, Event - Condition - Action (ECA) are used as a flexible tool to govern the relations between sensors and actuators. Runtime coordination and formal analysis becomes a necessity to avoid side effects mainly when applications are critical. In cooperation with CEA LETI/DACLE, we have worked on a case study for safe applications development in IoT and smart home environments [11].

[0] https://persyval-lab.org/en/sites/hpes

DAHU Project-Team

6. New Results

6.1. Highlights of the Year

Victor Vianu was elected member of Academia Europaea.
BEST PAPER AWARD :
[21] **Joint 25th International Conference on Rewriting Techniques and Applications and 12th International Conference on Typed Lambda Calculi and Applications**. S. SCHMITZ.

6.2. Distributed data management

Participants: Serge Abiteboul, Émilien Antoine, Victor Vianu.

The management of Web users' personal information is increasingly distributed across a broad array of applications and systems, including online social networks and cloud-based services. While users wish to share and integrate data using these systems, it is increasingly difficult to avoid the risks of unintended disclosures or unauthorized access by applications.

In [15], we study deduction in the presence of inconsistencies. Following previous works, we capture deduction via datalog programs and inconsistencies through violations of functional dependencies (FDs). We study and compare two semantics for datalog with FDs: the first, of a logical nature, is based on inferring facts one at a time, while never violating the FDs; the second, of an operational nature, consists in a fixpoint computation in which maximal sets of facts consistent with the FDs are inferred at each stage. Both semantics are nondeterministic, yielding sets of possible worlds. We introduce a PTIME (in the size of the extensional data) algorithm, that given a datalog program, a set of FDs and an input instance, produces a c-table representation of the set of possible worlds. Then, we propose to quantify nondeterminism with probabilities, by means of a probabilistic semantics. We consider the problem of capturing possible worlds along with their probabilities via probabilistic c-tables. We then study classical computational problems in this novel context. We consider the problems of computing the probabilities of answers, of identifying most likely supports for answers, and of determining the extensional facts that are most influential for deriving a particular fact. We show that the interplay of recursion and FDs leads to novel technical challenges in the context of these problems.

Jakub Kalas (ENS Cachan) spent 4 months in the team working on Personal Information Management Systems, using primarily positioning from data mobile phone and data from search engines.

6.3. Query Processing for the Web

Participants: Johann Brault-Baron, Arnaud Durand, Nadime Francis, Luc Segoufin, Cristina Sirangelo.

In many applications the output of a query may have a huge size and enumerating all the answers may already consume too many of the allowed resources. In this case it may be appropriate to first output a small subset of the answers and then, on demand, output a subsequent small numbers of answers and so on until all possible answers have been exhausted. To make this even more attractive it is preferable to be able to minimize the time necessary to output the first answers and, from a given set of answers, also minimize the time necessary to output the next set of answers - this second time interval is known as the *delay*. We have shown that this was doable with a almost linear preprocessing time and constant enumeration delay for first-order queries over structures of low degree [18]. We also presented a survey about this work at the Intl. Symp. on Theoretical Aspects of Computer Science (STACS) [22].

We have also been interested in querying data structured as graphs, which is nowadays spreading on the Web. Examples are social networks, linked data and the semantic web, via the RDF format. We have tackled the problem of answering queries over graph databases which are available only trough a given set of views. This is a common situation in many applications where access to data needs to be either controlled or optimized. In [19] we have studied when it is possible to rewrite over the views queries issued on the original data, and which query languages are needed for this purpose. We have considered views and queries expressed as Regular path queries, a very common graph query language.

6.4. Complexity in Counter Systems and Substructural Logics

Participant: Sylvain Schmitz.

The ties between propositional substructural logics (like linear logic, relevance logic, affine logic, etc.) on the one hand and extensions of vector addition systems on the other hand have long been known, as they lie for instance at the heart of undecidability proof of provability in linear logic. In a series a papers we recently revisited these connections with an eye on complexity issues. This allowed us to prove tight complexity bounds on provability in affine and contractive fragments of linear logic [20], in affine $(!, \oplus)$-Horn linear logic [16], and in implicational relevance logic [21] (an open problem for more than 25 years, with consequences on type inhabitation in the λI-calculus). Our work also yields a new Tower lower bound on reachability in branching vector addition systems [20], which entails the same lower bound for logics on XML trees [4], for which decidability is still open.

Although the connection with data logics might not seem obvious at first, the models of counter systems considered in these papers are tightly connected with logics for XML processing [5], [4]. Further investigations in the relationships between data logics, substructural logics, and counter systems are the main thrust behind the just accepted ANR PRODAQ project (see Section 8.1.1).

6.5. Incomplete Databases

Participants: Nadime Francis, Cristina Sirangelo.

Incomplete databases appear in several different scenarios. Intuitively, pieces of information might not be available, or can get lost due to failures in storage or transmission. Alternatively, some complex data managements tasks, such as data integration or data exchange, use incomplete databases as a model for databases with missing or unspecified information. In the context of the Web, these tasks have become even more crucial, which increased the need to handle incomplete databases. Given an incomplete database, one of the first question to answer is that of consistency: can we make sure that the incomplete database can be completed as a real database conforming to some specified schema.

Together with Claire David and Filip Murlak, we have considered this problem when incomplete instances are represented as incomplete XML documents, where labels and nodes might be missing, and we additionally assume the DOM semantics, meaning that nodes never lose their identity (otherwise, they are considered completely lost). These are further modeled as injective tree patterns using child and descendant relations. In [17], we close the question of the complexity of checking the consistency of such patterns with regards to a fixed regular tree language: it is polynomial for patterns that do not use child edges, and for patterns that use both, it is already NP-complete for patterns using at most two descendant edges per branch, the case for at most one descendant edge being already known to be polynomial.

In [12] we have studied the feasibility of query answering in the presence of incomplete information in data. In particular we have generalized conditions allowing classical query evaluation techniques to be applicable also in the presence of incompleteness. Our results show that conditions found in some of our previous work can be significantly relaxed so as to account for more complex semantics of incompleteness, originating in the fields of logic programming, programming semantics and data exchange.

DANTE Team

6. New Results

6.1. Highlights of the Year

6.1.1. The Internet of Things: A new equipments of excellence

Inaugurated last autumn, the very large scale IoT-LAB platform (https://www.iot-lab.info) is strengthening the capabilities of the FIT equipment of excellence dedicated to the Internet of Things. Offering a unique wide-ranging collection of equipment, these laboratories are available to both researchers and commercial companies alike.

IoT-LAB is a large-scale experimental platform for communicating objects and networks of sensors. It enables the rapid deployment of experiments and the collection of large amounts of data. It includes over 2700 sensor nodes, distributed over six sites in France, offering a wide range of different processor architectures and radio components. IoT-LAB is available for use on line. It is already used by over 300 users in forty countries, including around ten commercial companies. As of the end of October 2014, some 10 000 experiments had already been carried out.

6.1.2. Graph-based signal processing

Our first results towards the definition of a digital framework for signal processing on graphs constitutes an important outcome of DANTE's activity in 2014. Our participation to this emerging discipline was marked with several scientific recognitions: publication in the main DSP conference [14], involvement in the first ANR project focusing on this theme and retained for funding (2015-2019), we are in charge of the organisation of a Special Session dedicated to "Methodologies for signal processing on graphs" at Eusipco conference (2015).

6.1.3. Complex contagion process

Diffusion of innovation can be interpreted as a social spreading phenomena governed by the impact of media and social interactions. Although these mechanisms have been identified by quantitative theories, their role and relative importance are not entirely understood, since empirical verification has so far been hindered by the lack of appropriate data. Here we analyse a dataset recording the spreading dynamics of the world's largest Voice over Internet Protocol service to empirically support the assumptions behind models of social contagion. We show that the rate of spontaneous service adoption is constant, the probability of adoption via social influence is linearly proportional to the fraction of adopting neighbors, and the rate of service termination is time-invariant and independent of the behavior of peers. By implementing the detected diffusion mechanisms into a dynamical agent-based model, we are able to emulate the adoption dynamics of the service in several countries worldwide. This approach enables us to make medium-term predictions of service adoption and disclose dependencies between the dynamics of innovation spreading and the socioeconomic development of a country. This work was recently published in the Journal of the Royal Society Interface.

6.2. Diffusion and dynamic of complex networks

Participants: Márton Karsai [correspondant], Éric Fleury, Christophe Crespelle.

Time varying networks and the weakness of strong ties We analyse a mobile call dataset and find a simple statistical law that characterize the temporal evolution of users' egocentric networks. We encode this observation in a reinforcement process defining a time-varying network model that exhibits the emergence of strong and weak ties. We study the effect of time-varying and heterogeneous interactions on the classic rumor spreading model in both synthetic, and real-world networks. We observe that strong ties severely inhibit information diffusion by confining the spreading process among agents with recurrent communication patterns. This provides the counterintuitive evidence that strong ties may have a negative role in the spreading of information across networks.

Complex contagion process in spreading of online innovation [8]. Here we analyse a dataset recording the spreading dynamics of the world's largest Voice over Internet Protocol service to empirically support the assumptions behind models of social contagion. We show that the rate of spontaneous service adoption is constant, the probability of adoption via social influence is linearly proportional to the fraction of adopting neighbors, and the rate of service termination is time-invariant and independent of the behavior of peers. By implementing the detected diffusion mechanisms into a dynamical agent-based model, we are able to emulate the adoption dynamics of the service in several countries worldwide. This approach enables us to make medium-term predictions of service adoption and disclose dependencies between the dynamics of innovation spreading and the socio-economic development of a country.

The role of endogenous and exogenous mechanisms in the formation of R&D networks [10]. Here we propose a general modeling framework that includes both endogenous and exogenous mechanisms of in link formations in networks with tunable relative importance. The model contains additional ingredients derived from empirical observations, such as the heterogeneous propensity to form alliances and the presence of circles of influence, i.e. clusters of firms in the network. We test our model against the Thomson Reuters SDC Platinum dataset, one of the most complete datasets available nowadays, listing cross-country R&D alliances from 1984 to 2009. Interestingly, by fitting only three macroscopic properties of the network, this framework is able to reproduce a number of microscopic measures characterizing the network topology, including the distributions of degree, local clustering, path length and component size, and the emergence of network clusters. Furthermore, by estimating the link probabilities towards newcomers and established firms from the available data, we find that endogenous mechanisms are predominant over the exogenous ones in the network formation. This quantifies the importance of existing network structures in selecting partners for R&D alliances.

Controlling Contagion Processes in Time-Varying Networks [9]. In this project we derive an analytical framework for the study of control strategies specifically devised for time-varying networks. We consider the removal/immunization of individual nodes according the their activity in the network and develop a block variable mean-field approach that allows the derivation of the equations describing the evolution of the contagion process concurrently to the network dynamic. We derive the critical immunization threshold and assess the effectiveness of the control strategies. Finally, we validate the theoretical picture by simulating numerically the information spreading process and control strategies in both synthetic networks and a large-scale, real-world mobile telephone call dataset.

Data-driven spreading for the detection of weak ties [24]. In this work we propose a new method to infer the strength of social ties by using new data-driven simulation techniques. We qualify links by the importance they play during the propagation of information in the social structure. We apply data-driven spreading processes combined with a river-basin algorithmic method to identify links, which are the responsible to bring the information to large number of nodes. We investigate the correlations of the new importance measure with other conventional characteristics and identify their best combination through a percolation analysis to sophisticate further the assignment of social tie strengths. Finally we explore the role of the identified high importance links in control of globally spreading processes through data-driven SIR model simulations. These results point out that the size of infected population can be reduced considerably by weakening interactions through ties with high importance but zero overlap compared to strategies based on dyadic communications.

Dynamic Contact Network Analysis in Hospital Wards [18]. We analyse a huge and very precise trace of contact data collected during 6 months on the entire population of a rehabilitation hospital. We investigate the graph structure of the average daily contact network. Our main results are to unveil striking properties of this structure in the considered hospital, and to present a methodology that can be used for analyzing any dynamic complex network where nodes are classified into groups.

6.3. Performance analysis and networks protocols

Participants: Anthony Busson [correspondant], Thomas Begin, Isabelle Guérin Lassous.

Modeling and optimization of CSMA/CA in VANET [7]. We propose a simple theoretical model to compute the maximum spatial reuse feasible in a VANET. We focus on the ad hoc mode of the IEEE 802.11p standard. Our model offers simple and closed-form formulas on the maximum number of simultaneous transmitters, and on the distribution of the distance between them. It leads to an accurate upper bound on the maximum capacity. In order to validate our approach, results from the analytical models are compared to simulations performed with the network simulator NS-3. We take into account different traffic distributions (traffic of vehicles), and study the impact of this traffic on capacity. An application of this work is the parameterization of the CSMA/CA mechanism.

Fast and accurate approximate performance analysis of multi-server facilities [4]. Systems with multiple servers are common in many areas and their correct dimensioning is in general a difficult problem under realistic assumptions on the pattern of user arrivals and service time distribution. We present an approximate solution for the underlying $Ph/Ph/c/N$ queueing model. Our approximation decomposes the solution of the Ph/Ph/c/N queue into solutions of simpler $M/Ph/c/N$ and $Ph/M/c/N$ queues. To further mitigate dimensionality issues, for larger numbers of servers and/or service time phases, we use a reduced state approximation to solve the $M/Ph/c/N$ queue. The proposed approach is conceptually simple, easy to implement and produces generally accurate results for the mean number in the system, as well as the loss probability. Typical relative errors for these two quantities are below 5%. A very significant speed advantage compared to the numerical solution of the full $Ph/Ph/c/N$ queue can be gained as the number of phases representing the arrival process and/or the number of servers increases.

Interference and throughput in spectrum sensing cognitive radio networks using point processes . Spectrum sensing is vital for secondary unlicensed nodes to coexist and avoid interference with the primary licensed users in cognitive wireless networks. In this paper, we develop models for bounding interference levels from secondary network to the primary nodes within a spectrum sensing framework. Instead of classical stochastic approaches where Poisson point processes are used to model transmitters, we consider a more practical model which takes into account the medium access control regulations and where the secondary Poisson process is judiciously thinned in two phases to avoid interference with the secondary as well as the primary nodes. The resulting process will be a modified version of the Mate´rn point process. For this model, we obtain bounds for the complementary cumulative distribution function of interference and present simulation results which show the developed analytical bounds are quite tight. Moreover, we use these bounds to find the operation regions of the secondary network such that the interference constraint is satisfied on receiving primary nodes. We then obtain theoretical results on the primary and secondary throughputs and find the throughput limits under the interference constraint.

Modeling of IEEE 802.11 Multi-hop Wireless Chains with Hidden Nodes [11]. We follow up an existing modeling framework to analytically evaluate the performance of multi-hop flows along a wireless chain of four nodes. The proposed model accounts for a non-perfect physical layer, handles the hidden node problem, and is applicable under workload conditions ranging from flow(s) with low intensity to flow(s) causing the network to saturate. Its solution is easily and quickly obtained and delivers estimates for the expected throughput and for the datagram loss probability of the chain with a good accuracy.

Anticipation of ETX Metric to manage Mobility in Ad Hoc Wireless Networks [19]. When a node is moving in a wireless network, the routing metrics associated to its wireless links may reflect link quality degradations and help the routing process to adapt its routes. Unfortunately, an important delay between the metric estimation and its inclusion in the routing process makes this approach inefficient. In this paper, we introduce an algorithm that predicts metric values a few seconds in advance, in order to compensate the delay involved by the link quality measurement and their dissemination by the routing protocol. We consider classical metrics, in particular ETX (Expected Transmission Count) and ETT (Expected Transmission Time), but we combine their computations to our prediction algorithm. Extensive simulations show the route enhancement as the Packet Delivery Ratio (PDR) is close to 1 in presence of mobility.

6.4. Graphs & Signal Processing

Participants: Paulo Gonçalves [correspondant], Éric Fleury, Christophe Crespelle.

6.4.1. Signal Processing on Graphs

Semi-Supervised Learning for Graph to Signal Mapping: a Graph Signal Wiener Filter Interpretation [14]. We investigate a graph to signal mapping with the objective of analyzing intricate structural properties of graphs with tools borrowed from signal processing. We successfully use a graph-based semi-supervised learning approach to map nodes of a graph to signal amplitudes such that the resulting time series is smooth and the procedure efficient and scalable. Theoretical analysis of this method reveals that it essentially amounts to a linear graph-shift-invariant filter with the a priori knowledge put into the training set as input. Further analysis shows that we can interpret this filter as a Wiener filter on graphs. We finally build upon this interpretation to improve our results.

6.4.2. Graphs

(Nearly-)tight bounds on the contiguity and linearity of cographs [6]. In this paper we show that the contiguity and linearity of cographs on n vertices are both $O(\log n)$. Moreover, we show that this bound is tight for contiguity as there exists a family of cographs on n vertices whose contiguity is $\Omega(\log n)$. We also provide an $\Omega(\log n/\log\log n)$ lower bound on the maximum linearity of cographs on n vertices. As a by-product of our proofs, we obtain a min-max theorem, which is worth of interest in itself, stating equality between the rank of a tree and the minimum height of one of its path partitions.

6.4.3. Signal processing

Analysis of intrapartum foetal heart rate (FHR), enabling early detection of foetal acidosis to prevent asphyxia and labour adverse outcomes, remains a challenging signal processing task. In this direction, we carried out a series of works to characterize the fetal heart rate variability with specific attributes able to discriminate between healthy fetuses and fetuses presenting a risk of brain injury. Last year, we investigated two different approaches:

Nearest-Neighbor based Wavelet Entropy Rate Measures for Intrapartum Fetal Heart Rate Variability [23]. Firstly, we showed that a k-nearest neighbor procedure yields estimates for entropy rates that are robust and well-suited to FHR variability. Secondly, we experimentally proved that entropy rates measured on multiresolution wavelet coefficients permit to improve classification performance.

Impacts of labour first and second stages on Hurst parameter based intrapartum FHR analysis [22]. In this study, we proposed to quantify the FHR temporal dynamics with a Hurst exponent estimated within a wavelet framework. Analyses performed over a large (3049 records) and well documented database revealed that the evolution of the Hurst exponent during delivery, is significantly different for healthy fetuses and for acidotic fetuses.

6.5. Complex network metrology

Participant: Christophe Crespelle.

Measuring the Degree Distribution of Routers in the Core Internet [15]. Most current models of the internet rely on knowledge of the degree distribution of its core routers, which plays a key role for simulation purposes. In practice, this distribution is usually observed directly on maps known to be partial, biased and erroneous. This raises serious concerns on the true knowledge one may have of this key property. Here, we design an original measurement approach targeting reliable estimation of the degree distribution of core routers, without resorting to any map. It consists in sampling random core routers and precisely estimate their degree thanks to probes sent from many distributed monitors. We run and assess a large-scale measurement following this approach, carefully controlling and correcting bias and errors encountered in practice. The estimate we obtain is much more reliable than previous knowledge, and it shows that the true degree distribution is very different from all current assumptions.

Measuring Routing Tables in the Internet [21]. The most basic function of an Internet router is to decide, for a given packet, which of its interfaces it will use to forward it to its next hop. To do so, routers maintain a routing table, in which they look up for a prefix of the destination address. The routing table associates an interface of the router to this prefix, and this interface is used to forward the packet. We explore here a new measurement method based upon distributed UDP probing to estimate this routing table for Internet routers.

DEDUCTEAM Exploratory Action

6. New Results

6.1. Highlights of the Year

In the framework of the *BWare* project, Pierre Halmagrand, David Delahaye, Damien Doligez, and Olivier Hermant designed a new version of the *B* set theory using deduction modulo, in order to automatically verify a large part of the proof obligations of the benchmark of *BWare*, which consists of proof obligations coming from the modeling of industrial applications (about 13,000 proof obligations). Using this *B* set theory modulo with *Zenon Modulo*, as well as some other extensions of *Zenon*, such as typed proof search and arithmetic (implemented by Guillaume Bury), we are able to automatically verify more than 95% of the proof obligations of *BWare*, while the regular version of *Zenon* is only able to prove less than 1% of these proof obligations. This is a real breakthrough for the *BWare* project, but also for automated deduction in general, as it tends to show that deduction modulo is the way to go when reasoning modulo theories.

6.2. Termination

Frédéric Blanqui, together with Jean-Pierre Jouannaud (Univ. Paris 11) and Albert Rubio (Technical University of Catalonia), have finished their work on a new version of the higher-order recursive path ordering (HORPO) [44], [43], a decidable monotone well-founded relation that can be used for proving the termination of higher-order rewrite systems by checking that rules are included in it. This new version, called the computability path ordering (CPO), appears to be the ultimate improvement of HORPO in the sense that this definition captures the essence of computability arguments *à la* Tait and Girard [37], therefore explaining the name of the improved ordering. It has been shown that CPO allows to consider higher-order rewrite rules in a simple type discipline with inductive types, that most of the guards present in the recursive calls of its core definition cannot be relaxed in any natural way without losing well-foundedness, and that the precedence on function symbols cannot be made more liberal anymore. This new result is described in a 41-pages papers available on Frédéric Blanqui's web page which has been submitted to a journal for publication. A Prolog implementation of CPO is also available on Albert Rubio's web page.

Frédéric Blanqui revised his work on the compatibility of Tait and Girard's notion of computability for proving the termination of higher-order rewrite systems when matching is done modulo $\beta\eta$-equivalence. In particular, he showed that computability is preserved by leaf-β-expansion, a key property for dealing with higher-order pattern-matching. This work is described in a 46-pages paper available on his web page which has been submitted to a journal for publication.

Frédéric Blanqui did some historical investigations on fixpoint theorems in posets used for instance for defining the semantics of non-basic inductive types (i.e. types with constructors taking functions as arguments) and the termination of functions defined by induction on such non-basic inductive types. These theorems assume the function either extensive or monotone. However, as shown by Salinas in [48], these two conditions can be subsumed by a more general one. Frédéric Blanqui slightly improved this condition further by using results by Hartogs, Rubin and Rubin, and Abian and Brown. This work is described in a 10-pages note available on his web page [20].

Kim Quyen Ly finished the development of a new version faster, safer (proved correct in Coq) and standalone version of Rainbow, based on Coq extraction mechanism. She defended her PhD thesis [11] on the automated verification of termination certificates in October.

6.3. Proof and type theory modulo rewriting

Ali Assaf defined a sound and complete embedding of the cumulative universe hierarchy of the *calculus of inductive constructions* (CIC) in the $\lambda\Pi$-calculus modulo rewriting [18]. By reformulating universes in the Tarski style, he showed that we can make cumulativity explicit without losing any typing power. This result refines the translation used by Coqine, which was unsound because it collapsed the universe hierarchy to a single type universe. It also sheds some light on the metatheory of Coq and its connection to Martin-Löf's intuitionistic type theory. This work was presented at the TYPES meeting in Paris.

Frédéric Gilbert and Olivier Hermant defined new encodings from classical to intuitionistic first-order logic. These encodings, based on the introduction of double negations in formulas, are tuned to satisfy two purposes jointly: basing their specifications on the definition of *classical connectives* inside intuitionistic logic – which is the property of *morphisms*, and reducing their impact on the shape and size of formulas, by limiting as much as possible the number of negations introduced. This paper has been submitted.

Raphael Cauderlier and Catherine Dubois defined a shallow embedding of an object calculus (formalized by Abadi and Cardelli), in the $\lambda\Pi$-calculus modulo rewriting . The main result concerns the encoding of subtyping. This encoding shows that rewriting is an effective help for handling of subtyping proofs. The implementation in Dedukti, Sigmaid. This work has been presented at the TYPES 2014 meeting in Paris. A paper has been submitted.

Ali Assaf, Olivier Hermant and Ronan Saillard defined a rewrite system such that all strongly normalizable proof term can be typed in Natural Deduction modulo this rewrite system. This work is inspired by Statman's work [49], and can be understood as an encoding of intersection types.

Guillaume Burel showed how to get rewriting systems that admit cut by using standard saturation techniques from automated theorem proving, namely ordered resolution with selection, and superposition. This work relies on a view of proposition rewriting rules as oriented clauses, like term rewriting rules can be seen as oriented equations. This also lead to introduce an extension of deduction modulo with *conditional* term rewriting rules. This work was presented at the RTA-TLCA conference in Vienna [15].

Gilles Dowek, has generalized the notion of super-consistency to the lambda-Pi-calculus modulo theory and proved this way the termination of the embedding of various formulations of Simple Type Theory and of the Calculus of Constructions in the Lambda-Pi calculs modulo theory.

Gilles Dowek and Alejandro Díaz-Caro have finished their work on the extension of Simply Typed Lambda-Calculus with Type Isomorphisms. This work has been presented at the Types meeting and recently accepted for publication in the Theoretical Computer Science journal [26].

Gilles Dowek and Ying Jiang have given a new proof of the decidability of reachability in alternating pushdown systems, based on a cut-elimination theorem.

Vaston Costa presented to the group a new structure to represent proofs through references rather than copy. The structure, called Mimp-graph, was initially developed for minimal propositional logic but the results have been extended to first-order logic. Mimp-graph preserves the ability to represent any Natural Deduction proof and its minimal formula representation is a key feature of the mimp-graph structure, it is easy to distinguish maximal formulas and an upper bound in the length of the reduction sequence to obtain a normal proof. Thus a normalization theorem can be proved by counting the number of maximal formulas in the original derivation. The strong normalization follow as a direct consequence of such normalization, since that any reduction decreases the corresponding measures of derivation complexity. Sharing for inference rules is performed during the process of construction of the graph. This feature is very important, since we intend to use this graph in automatic theorem provers.

6.4. Automated theorem proving

Guillaume Bury defined a sound and complete extension of the tableaux method to handle linear arithmetic. The rules are based on a variant of the simplex algorithm for rational and real linear arithmetic, and a Branch&Bound algorithm for integer arithmetic.

Guillaume Bury defined an encoding of analytical tableaux rules as a theory for smt solvers. The theory acts like a lazy cnf conversion during the proof search and allows to integrate the cnf conversion into the resolution proof for unsatisfiable formulas. This work was implemented in mSAT.

Simon Cruanes added many improvements to Logtk, in particular a better algorithm to reduce formulas to Clausal Normal Form. A presentation of its design and implementation has been made at PAAR 2014[16]. He also used Zipperposition as a testbed for integer linear arithmetic; a sophisticated inference system for this fragment of arithmetic was designed and implemented in Zipperposition, including many redundancy criteria and simplification rules that make it efficient in practice. The arithmetic-enabled Zipperposition version entered CASC-J7, the annual competition of Automated Theorem Provers, in the first-order theorems with linear arithmetic division where it had very promising results (on integer problems only, since Zipperposition doesn't handle rationals).

Another extension of Zipperposition has been performed by Julien Rateau, Simon Cruanes, and David Delahaye, in order to deal with a fragment of set theory in the same vein as the *STR*$\dot{+}$*VE*$\subseteq$ prover [40]. This extension relies on a specific normal form of literal, which only involves the $\subseteq$, $\cap$, $\cup$, and complement set operators. In the future, the idea is to use this extension in the framework of the *BWare* project to verify *B* proof obligations coming from industrial benchmarks.

The current effort of research on Zipperposition focuses on extending superposition to handle structural induction, following the work from [45]. The current prototype is able to prove simple properties on natural numbers, binary trees and lists.

Kailiang Ji defined a set of rewrite rules for the equivalence between CTL formulas (denote them as R_{CTL}), by taking them as terms of designed predicates. For a given transition system model, we transform it into a set of rewrite rules (denote them as R_m). Then any CTL property of the transition system can be proved in deduction modulo $R_{CTL} \cup R_m$, by specifying the model checking problems into designed first-order formulas. This method was implemented in iProver Modulo, and the experimental evaluation was reported in workshop of Locali 2014.

6.5. Algebraic λ-calculus

Ali Assaf, Alejandro Díaz-Caro, Simon Perdrix, Christine Tasson, and Benoit Valiron completed a journal paper covering results on different algebraic extensions of the λ-calculus [12]. These extensions equip the calculus with an additive and a scalar-multiplicative structure, and their set of terms is closed under linear combinations. Two such extensions, the *algebraic λ-calculus* and the *linear-algebraic λ-calculus* arise independently in different contexts – the former is a fragment of the differential λ-calculus, the latter is a candidate λ-calculus for quantum computation – and have different operational semantics. In this paper, the authors showed how the two approaches relate to each other. They showed that the the first calculus follows a call-by-name strategy while the second follows a call-by-value strategy. They proved that the two can simulate each other using algebraic extensions of *continuation passing style* (CPS) translations that are sound and complete.

DEFI Project-Team

6. New Results

6.1. Qualitative methods for inverse scattering problems

6.1.1. Identifying defects in an unknown background using differential measurements

Participants: Lorenzo Audibert, Houssem Haddar.

With Alexandre Girard, we developed a new qualitative imaging method capable of selecting defects in complex and unknown background from differential measurements of farfield operators: i.e. far measurements of scattered waves in the cases with and without defects. Indeed, the main difficulty is that the background physical properties are unknown. Our approach is based on a new exact characterization of a scatterer domain in terms of the far field operator range and the link with solutions to so-called interior transmission problems. We present the theoretical foundations of the method and some validating numerical experiments in a two dimensional setting [10]. This work is based on the generalized formulation of the Linear Sampling Method with exact characterization of targets in terms of farfield measurements that has been introduced in [1].

6.1.2. The Factorization Method for a Cavity in an Inhomogeneous Medium

Participants: Houssem Haddar, Shixu Meng.

With F. Cakoni we considered the inverse scattering problem for a cavity that is bounded by a penetrable anisotropic inhomogeneous medium of compact support where on is interested in determining the shape of the cavity from internal measurements on a curve or surface inside the cavity. We derived a factorization method which provides a rigorous characterization of the support of the cavity in terms of the range of an operator which is computable from the measured data. The support of the cavity is determined without a-priori knowledge of the constitutive parameters of the surrounding anisotropic medium provided they satisfy appropriate physical as well as mathematical assumptions imposed by our analysis. Numerical examples were given showing the viability of our method [7].

6.1.3. Asymptotic analysis of the transmission eigenvalue problem for a Dirichlet obstacle coated by a thin layer of non-absorbing media

Participant: Houssem Haddar.

With F. Cakoni and N. Chaulet we considered the transmission eigenvalue problem for an impenetrable obstacle with Dirichlet boundary condition surrounded by a thin layer of non-absorbing inhomogeneous material. We derived a rigorous asymptotic expansion for the first transmission eigenvalue with respect to the thickness of the thin layer. Our convergence analysis is based on a Max–Min principle and an iterative approach which involves estimates on the corresponding eigenfunctions. We provided explicit expressions for the terms in the asymptotic expansion up to order 3 [3].

6.1.4. Boundary Integral Equations for the Transmission Eigenvalue Problem for Maxwell's Equations

Participants: Houssem Haddar, Shixu Meng.

In this work, we considered the transmission eigenvalue problem for Maxwell's equations corresponding to non-magnetic inhomogeneities with contrast in electric permittivity that changes sign inside its support. Following the approach developed by Cossonnière-Haddar in the scalar case, we formulate the transmission eigenvalue problem as an equivalent homogeneous system of boundary integral equation and prove that assuming that the contrast is constant near the boundary of the support of the inhomogeneity, the operator associated with this system is Fredholm of index zero and depends analytically on the wave number. Then we show the existence of wave numbers that are not transmission eigenvalues which by an application of the analytic Fredholm theory implies that the set of transmission eigenvalues is discrete with positive infinity as the only accumulation point. This is a joint work with F. Cakoni.

6.1.5. Invisibility in scattering theory

Participant: Lucas Chesnel.

We investigated a time harmonic acoustic scattering problem by a penetrable inclusion with compact support embedded in the free space. We considered cases where an observer can produce incident plane waves and measure the far field pattern of the resulting scattered field only in a finite set of directions. In this context, we say that a wavenumber is a non-scattering wavenumber if the associated relative scattering matrix has a non trivial kernel. Under certain assumptions on the physical coefficients of the inclusion, we showed that the non-scattering wavenumbers form a (possibly empty) discrete set. This result is important in the justification of certain reconstruction techniques like the Linear Sampling Method in practical applications.

In a second step, for a given real wavenumber and a given domain D, we developed a constructive technique to prove that there exist inclusions supported in D for which the corresponding relative scattering matrix is null. These inclusions have the important property to be impossible to detect from far field measurements. The approach leads to a numerical algorithm which allows to provide examples of (approximated) invisible inclusions. This is a joint work with A.-S. Bonnet-Ben Dhia and S.A. Nazarov [11].

6.1.6. Invisibility in electrical impedance tomography

Participant: Lucas Chesnel.

We adapted the technique to construct invisible isotropic conductivities in for the point electrode model in electrical impedance tomography. Again, the theoretical approach, based on solving a fixed point problem, is constructive and allows the implementation of an algorithm for approximating the invisible perturbations. We demonstrated the functionality of the method via numerical examples. This a joint work with N. Hyvönen and S. Staboulis [13].

6.1.7. A quasi-backscattering problem for inverse acoustic scattering in the Born regime

Participants: Houssem Haddar, Jacob Rezac.

In this work we propose a data collection geometry in which to frame the inverse scattering problem of locating unknown obstacles from far-field measurements of time-harmonic scattering data. The measurement geometry, which we call the quasi-backscattering set-up, is configurated such that one device acts as a transmitter and a line of receivers extends in one-dimension a small distance from the transmitter. We demonstrate that the data collected can be used to locate inhomogeneities whose physical properties are such that the Born approximation applies. In particular, we are able to image a two-dimensional projection of the location of an obstacle by checking if a test function which corresponds to a point in $\mathbb{R}^2$ belongs to the range of a measurable operator. The reconstruction algorithm is based on the MUSIC (Multiple SIgnal Classification) algorithm.

6.2. Iterative Methods for Non-linear Inverse Problems

6.2.1. Inverse medium problem for axisymmetric eddy current models

Participants: Houssem Haddar, Zixian Jiang, Mohamed Kamel Riahi.

We continued our developments of shape optimization methods for inclusion detection in an axisymmetric eddy current model. This problem is motivated by non-destructive testing methodologies for steam generators. We finalized a joint work with A. Lechleiter on numerical methods for the solution of the direct problem in weighted Sobolev spaces using approriate Dirichlet-to-Neumann mappings to bound the computational domain. We are also finalized jointly with M. El Guedri the work on inverse solver using a regularized steepest descent method for the problem of identifying a magnetite deposits using axial eddy current probe. We addressed two issues:

- We developed asymptotic models to identify thin highly conducting deposits. We derived three possible asymptotic models that can be exploited in the inverse problem. We are about to finalize a preprint on this topic.
- We extended the inverse scheeme to 3D configurations with axisymmetry at infinity: this includes exact charactherization of the shape derivative for a mixed formulation of eddy current problems and a parametric inversion scheme based on a pre-defined discrete grid for deposit location [14].

6.2.2. The conformal mapping method and free boundary problems

Participant: Houssem Haddar.

Together with R. Kress we employed a conformal mapping technique for the inverse problem to reconstruct a perfectly conducting inclusion in a homogeneous background medium from Cauchy data for electrostatic imaging, that is, for solving an inverse boundary value problem for the Laplace equation. In a recent work we proposed an extension of this approach to inverse obstacle scattering for time-harmonic waves, that is, to the solution of an inverse boundary value problem for the Helmholtz equation. The main idea is to use the conformal mapping algorithm in an iterative procedure to obtain Cauchy data for a Laplace problem from the given Cauchy data for the Helmholtz problem. We presented the foundations of the method together with a convergence result and exhibit the feasibility of the method via numerical examples. We are currently investigating the extention of this method to solve free boundary value problems.

6.2.3. A steepest descent method for inverse electromagnetic scattering problems

Participant: Houssem Haddar.

Together with N. Chaulet, we proposed the application of a non linear optimization techniques to solve the inverse scattering problems for the $3D$ Maxwell's equations with generalized impedance boundary conditions. We characterized the shape derivative in the case where the GIBC is defined by a second order surface operator. We then applied a boundary variation method based on a regularized steepest descent to solve the 3-D inverse problem with partial farfield data. The obtained numerical results demonstrated the possibility of identifying the shape of coated objects as well as the parameters of the coating in the $3D$ Maxwell case [4].

6.2.4. A posteriori error estimates: Application to Electrical Impedance Tomography

Participants: Olivier Pantz, Matteo Giacomini.

One of the main problem in shape optimization problems is due to the fact that the gradient is never computed exactly. When the current solution is far from a local optimum, this is not a problem: even a rough approximation of the gradient enable us to exhibit a descent direction. On the contrary, when close to a local optimal, a very precise computation of the gradient is needed. Together with Karim Trabelsi, we propose to use a-posteriori error estimates to evaluate the error made on the computation of the gradient. This enables us to ensure that at each step, a genuine descent direction is used in the gradient method. Our method has been applied to the minimization of the Kohn-Vogelius functional in the context of electrical impedance tomography. An article is currently in preparation.

6.2.5. A robust stopping rule for EM algorithm with applications to SAXS measurements

Participants: Federico Benvenuto, Houssem Haddar.

The aim of this work was to develop a fully automatic method for the reconstruction of the volume distribution of diluted polydisperse non-interacting nanoparticles with identical shape from Small Angle X-ray Scattering measurements. The described method solves a maximum likelihood problem with a positivity constraint on the solution by means of an Expectation Maximization iterative scheme coupled with an innovative type of regularization. Such a regularization, together with the positivity constraint results in high fidelity quantitative reconstructions of particle volume distributions making the method particularly effective in real applications. The performance of the method on synthetic data in the case of uni- and bi-modal particle volume distributions are shown. Moreover, the reliability of the method is tested when applied to real data provided by a Xenocs device prototype. Finally, the method can be extended to the analysis of the particle distribution for different types of nano-structures.

6.3. Shape and topology optimization

6.3.1. Stacking sequence and shape optimization of laminated composite plates

Participant: Grégoire Allaire.

We consider the optimal design of composite laminates by allowing a variable stacking sequence and in-plane shape of each ply. In order to optimize both variables we rely on a decomposition technique which aggregates the constraints into one unique constraint margin function. Thanks to this approach, a rigorous equivalent bi-level optimization problem is established. This problem is made up of an inner level represented by the combinatorial optimization of the stacking sequence and an outer level represented by the topology and geometry optimization of each ply. We propose for the stacking sequence optimization an outer approximation method which iteratively solves a set of mixed integer linear problems associated to the evaluation of the constraint margin function. For the topology optimization of each ply, we lean on the level set method for the description of the interfaces and the Hadamard method for boundary variations by means of the computation of the shape gradient. An aeronautic test case is exhibited subject to different constraints, namely compliance, reserve factor and first buckling load. This is joint work with G. Delgado.

6.3.2. Thickness control in structural optimization via a level set method

Participant: Grégoire Allaire.

In the context of structural optimization via a level-set method we propose a framework to handle geometric constraints related to a notion of local thickness. The local thickness is calculated using the signed distance function to the shape. We formulate global constraints using integral functionals and compute their shape derivatives. We discuss different strategies and possible approximations to handle the geometric constraints. We implement our approach in two and three space dimensions for a model of linearized elasticity. As can be expected, the resulting optimized shapes are strongly dependent on the initial guesses and on the specific treatment of the constraints since, in particular, some topological changes may be prevented by those constraints. This is a joint work with G. Michailidis

6.4. Asymptotic Analysis

6.4.1. Effective boundary conditions for thin periodic coatings

Participants: Mathieu Chamaillard, Houssem Haddar.

This topic is the object of a collaboration with Patrick Joly and is a continuation of our earlier work on interface conditions done in the framework of the PhD thesis of Berangère Delourme. The goal here is to derive effective conditions that model scattering from thin periodic coatings where the thickness and the periodicity are of the same length but very small compared to the wavelength. The originality of our work, compared to abundant literature is to consider the case of arbitrary geometry (2-D or 3-D) and to consider higher order approximate models. We formally derived third order effective conditions after exhibiting the full asymptotic expansion of the solution in terms of the periodicity length.

6.4.2. Homogenization of the transmission eigenvalue problem with applications to inverse problems

Participant: Houssem Haddar.

In a joint work with F. Cakoni and I. Harris, we consider the interior transmission problem associated with the scattering by an inhomogeneous (possibly anisotropic) highly oscillating peri-odic media. We show that, under appropriate assumptions, the solution of the interior transmission problem converges to the solution of a homogenized problem as the period goes to zero. Furthermore, we prove that the associ-ated real transmission eigenvalues converge to transmission eigenvalues of the homogenized problem. Finally we show how to use the first transmission ei-genvalue of the period media, which is measurable from the scattering data, to obtain information about constant effective material properties of the periodic media. The obtained convergence results are not optimal. Such results with rate of convergence involve the analysis of the boundary correction and will be subject of a forthcoming paper.

6.4.3. Homogenization of electrokinetic models in porous media

Participant: Grégoire Allaire.

With R. Brizzi, J.-F. Dufrêche, A. Mikelic and A. Piatnitski, are interested in the homogenization (or upscaling) of a system of partial differential equations describing the non-ideal transport of a N-component electrolyte in a dilute Newtonian solvent through a rigid porous medium. Our recent work has focused on the so-called non-ideal case. Namely we consider the mean spherical approximation (MSA) model which takes into account finite size ions and screening effects. On the one hand we established a rigorous homogenized transport model starting from this microscopic model. On the other hand we did numerical simulations to ompute the corresponding effective parameters and make systematic comparisons between the idea model and the MSA model.

6.4.4. Modeling and Simulation of the Mechanical behavior of Vesicles and Red Blood Cells

Participant: Olivier Pantz.

6.4.4.1. Highly anisotropic thin shells

With K. Trabelsi (IPSA), we have proposed a new justification of various non linear highly anisotropic elastic shell models. Among others, we do derive the so called Helfrich functional, that describe the behavior of the lipid bilayer of the vesicle and red blood cells. Our results will soon be published in MEMOCS (Mathematics and Mechanics Complex Systems).

6.4.4.2. Minimization of the Helfrich functional

Our work with K. Trabelsi established that the mechanical behavior of vesicles and red blood cells can be approximated by thin non linear anisotropic elastic shells. Minimizing directly the Helfrich functional is not an easy task from the numerical point of view. Most methods require the use of high order finite elements and stabilization techniques so to prevent mesh degeneration. Instead, we propose to approximate the two dimensional membrane of a vesicle (or red blood cell) by a three dimensional non linear elastic body of small thickness. Firstly, this enable us to use standard finite elements and discretization (basically Lagrange of degree 2). Secondly, the discretized formulation is intrinsically stable, so no stabilization is needed. Finally, even if it leads us to solve a three dimensional problem (instead of the two dimensional initial one), it is no more costly than a direct two dimensional approach as the scale of the mesh can be chosen to be of the same order than the "thickness" of the shell. We have already obtained encouraging results for vesicles. We plan to extend them to the case of vesicles with spontaneous curvature and to red blood cells. Moreover, we are considering different strategies to minimize the computational cost (that is already quite satisfying compared with some other methods).

6.4.5. Modeling of Damage and Fracture

Participant: Olivier Pantz.

6.4.5.1. Fracture as limit of Damage

With Leila Azem (PhD Student), we use a model introduced by G. Allaire, F. Jouve and N. Van Goethem in a previous work to simulate the propagation of fracture. The main idea is to approximate the fracture as a damage material and to compute the evolution of the path of the crack using a shape gradient analysis. Our main contribution consists to propose to use a material derivative approach to compute the shape gradient. The advantage is that it drastically simplify the evaluation of the shape gradient, as no regularization is needed and no jump terms as to be computed on the interface bewteen the healthy and damaged areras. An article is currently in preparation to present our results.

6.4.5.2. Fracture with penalization of the jump

With Leila Azem, we propose to approximate a model of fracture with penalization of the jump of the displacement as a limit of a damage model. This is achieved by a specific choice of the softness of the damage material with respect to the cost to turn material from a healthy to a damaged state. We have carried out a formal analysis to justify our approach and have already obtained several numerical results.

6.5. Diffusion MRI

Participants: Jing-Rebecca Li, Houssem Haddar, Simona Schiavi, Khieu Van Nguyen, Gabrielle Fournet, Dang Van Nguyen.

Diffusion Magnetic Resonance Imaging (DMRI) is a promising tool to obtain useful information on microscopic structure and has been extensively applied to biological tissues. In particular, we would like to focus on two applications:

- Inferring from DMRI measurements changes in the cellular volume fraction occurring upon various physiological or pathological conditions. This application is one of the first to show the promise of DMRI because it can detect acute cerebral ischemia (cell swelling) on the basis of lower than normal apparent diffusion coefficient a few minutes after stroke;
- Estimating the average cell size in the case of tumor imaging This application is useful as a diagnostic tool as well as a tool for the evaluation of tumor treatments;

For both of the above applications we approach the problem via the following steps:

- Construct reduced models of the multiple-compartment Bloch-Torrey partial differential equation (PDE) using homogenization methods.
- Invert the resulting reduced models for the biological parameters of interest: the cellular volume fraction in the first case, and the average distance between neighboring cells in the second case.

We obtained the following results.

- We derived using homogenization techniques an asymptotic model of the diffusion MRI signal for finite pulse magnetic field gradient sequences in the long diffusion time regime and numerically verified it using a Finite Element method for both isotropic and anisotropic diffusion configurations in three dimensions. This resulted in 2 publications.
- We derived a new asymptotic model of the diffusion MRI signal for low gradient strengths that is valid for a wide range of diffusion time scales. An article describing our results is under preparation.
- We performed a numerical study of a cylinder model of the diffusion MRI signal for neuronal dendrite trees. This resulted in 1 publication.
- We implemented a compressed sensing method for obtaining T2-weighted images in shorter scanning time and this method was used to segment nerve cells of the Aplysia Californica at the MRI center Neurospin. An article describing our results is under preparation.
- We participated in the characterization of glioma microcirculation and tissue features in a rat brain model using diffusion-encoding magnetic field gradient pulses sequences, working along with collaborators at the high field brain MRI center Neurospin. This resulted in 1 publication.
- We performed Monte-Carlo simulation of blood flow in micro-vessels in the brain with the goal of using the results to explain the MRI signal drop due to incoherent flow in the micro-vessels. This is an ongoing project.

DEMAR Project-Team

5. New Results

5.1. Modelling and identification of the sensory-motor system

5.1.1. Whole Body Center of Mass Estimation with Portable Sensors: Using the Statically Equivalent Serial Chain and a Kinect

Participants: Alejandro Gonzalez de Alba, Mitsuhiro Hayashibe, Vincent Bonnet, Philippe Fraisse.

The trajectory of the whole body center of mass (CoM) is useful as a reliable metric of postural stability. If the evaluation of a subject-specific CoM were available outside of the laboratory environment, it would improve the assessment of the effects of physical rehabilitation. A method is developed tot enable tracking CoM position using low-cost sensors such that it can be moved around by a therapist or easily installed inside a patientâs home. We compare the accuracy of a personalized CoM estimation using the statically equivalent serial chain (SESC) method and measurements obtained with the Kinect to the case of a SESC obtained with high-end equipment (Vicon). We also compare these estimates to literature-based ones for both sensors. The method was validated with seven able-bodied volunteers for whom the SESC was identified using 40 static postures. The literature-based estimation with Vicon measurements had a average error 24.9 $\pm$ 3.7 mm; this error was reduced to 12.8 $\pm$ 9.1 mm with the SESC identification. When using Kinect measurements, the literature-based estimate had an error of 118.4 $\pm$ 50.0 mm, while the SESC error was 26.6 $\pm$ 6.0 mm. The subject-specific SESC estimate using low-cost sensors has an equivalent performance as the literature-based one with high-end sensors. The SESC method can improve CoM estimation of elderly and neurologically impaired subjects by considering variations in their mass distribution.

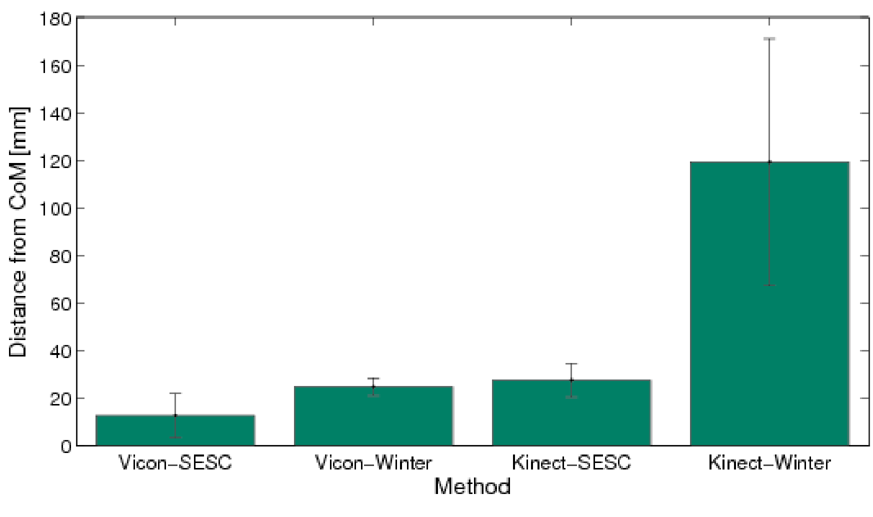


Figure 2. The bars correspond to the averaged results for all seven subjects. We observe an increase in the accuracy of the identified SESCs with respect to the literature estimates; we found that the performance of the Kinect-SESC is equivalent to that of the literature-based estimate using high-end sensors.

5.1.2. A Personalized Balance Measurement for Home-based Rehabilitation

Participants: Alejandro Gonzalez de Alba, Philippe Fraisse, Mitsuhiro Hayashibe.

We use the subject-specific center of mass (CoM) estimate offered by the statically equivalent serial chain (SESC) method and the zero rate of change of angular momentum (ZRAM) concept to evaluate balance for a series of dynamic motions. Two healthy subjects were asked to stand on a Wii balance board and their SESC parameters were identified. A set of dynamic motions was to evaluate the rate of change of centroidal angular momentum and the distance of the ZRAM point to the center line of the support polygon. We found a good match between both balance metrics. As an application example, the subjects performed a dynamic motion (such as walking and abruptly stopping) and the stability was determined in real-time using the ZRAM point from the personalized CoM trajectory. This was implemented with a real-time balance visualization tool based on Kinect measurements for home-based rehabilitation.

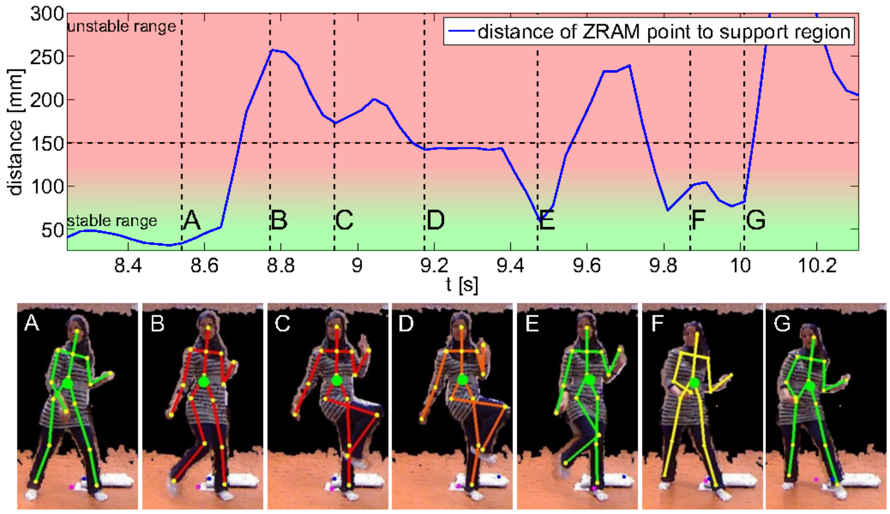


Figure 3. The distance between the ZRAM point and the center of the support polygon can be used to determine balance in real-time and shown as feedback during a physical rehabilitation program. In this example, the subject's skeleton is colored in red to indicate an unstable movement and in green for a stable one.

5.1.3. Methodology for automatic movement cycle extraction using Switching Linear Dynamic System

Participants: Roberto de Souza Baptista, Mitsuhiro Hayashibe, Antônio P. L. Bó [Univ. Brasilia].

Human motion assessment is key for motor-control rehabilitation. Using standardized definifions and spatiotemporal features - usually presented as a movement cycle diagram- specialists can associate kinematic measures to progress in rehabilitation therapy or motor impairment due to trauma or disease. Although devices for capturing human motion today are cheap and widespread, the automatic interpretation of kinematic data for rehabilitation is still poor in terms of quantitative performance evaluation. In this paper we present an automatic approach to extract spatiotemporal features from kinematic data and present it as a cycle diagram. This is done by translating standard definitions from human movement analisys into mathematical elements of a Switching Linear Dynamic System model. The result is a straight-forward procedure to learn a tracking model from a sample execution. This model is robust when used to automatically extract the movement cycle diagram of the same motion (the Sit-Stand-Sit, as an example) executed in different subject-specific manner such as his own motion speed.

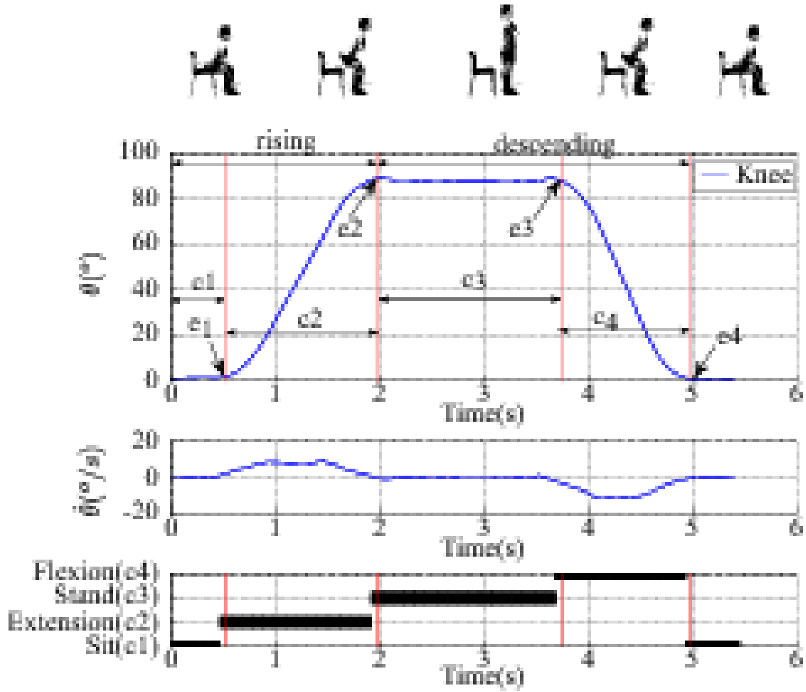


Figure 4. Training dataset consisting of one execution of the Sit-Stand-Sit movement cycle. Events (e_i), components (c_i) and the rising and descending phases are marked. θ and $\dot{\theta}$ indicates angle and angular velocity.

5.1.4. Real-time Muscle Deformation via Decoupled Modeling of Solid and Muscle Fiber Mechanics

Participants: Yacine Berranen, Mitsuhiro Hayashibe, David Guiraud, Benjamin Gilles.

This work presents a novel approach for simulating 3D muscle deformations with complex architectures. The approach consists in choosing the best model formulation in terms of computation cost and accuracy, that mixes a volumetric tissue model based on finite element method (3D FEM), a muscle fiber model (Hill contractile 1D element) and a membrane model accounting for aponeurosis tissue (2D FEM). The separate models are mechanically binded using barycentric embeddings. Our approach allows the computation of several fiber directions in one coarse finite element, and thus, strongly decreases the required finite element resolution to predict muscle deformation during contraction. Using surface registration, fibers tracks of specific architecture can be transferred from a template to subject morphology, and then simulated. As a case study, three different architectures are simulated and compared to their equivalent one dimensional Hill wire model simulations.

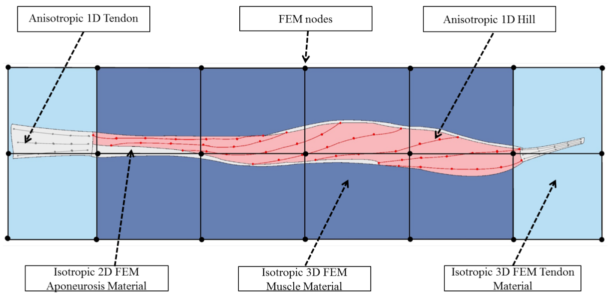


Figure 5. Muscle multi-model scheme: The different models are linked via barycentric embeddings. This approach strongly decreases the required finite element resolution to predict muscle deformation during contraction

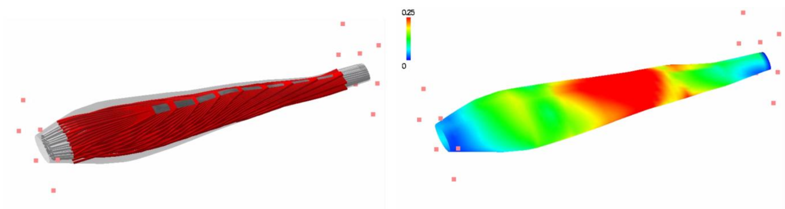


Figure 6. Left: Hybrid model of bipennate muscle, right: internal stress due to isometric contraction

5.1.5. Adaptive model for viscoelastic solids

Participants: Benjamin Gilles, Maxime Tournier, Matthieu Nesme, Francois Faure.

A new adaptive model for viscoelastic solids is presented in [50], [34]. Unlike previous approaches, it allows seamless transitions, and simplifications in deformed states. The deformation field is generated by a set of physically animated frames. Starting from a fine set of frames and mechanical energy integration points, the model can be coarsened by attaching frames to others, and merging integration points. Since frames can be attached in arbitrary relative positions, simplifications can occur seamlessly in deformed states, without returning to the original shape, which can be recovered later after refinement. We propose a new class of velocity-based simplification criterion based on relative velocities. Integration points can be merged to reduce the computation time even more, and we show how to maintain constant elastic forces through the levels of detail. This meshless adaptivity allows significant improvements of computation time.

5.1.6. Functional Brain Stimulation: Filling the gap between micro- and direct electrical stimulation of the brain in order to better understand and innovate

Participants: Marion Vincent, Olivier Rossel, Mitsuhiro Hayashibe, Guillaume Herbet, Hugues Duffau [Neurosurgery Department, CHU, Montpellier], David Guiraud, François Bonnetblanc.

Micro-stimulation (MES) and Direct electrical stimulation (DES) of the brain are both used to perform in vivo functional mapping of the brain in fundamental neuroscience and neurosurgery respectively. The former is performed in animal experiments while the latter is performed on humans in the operative room. Very recently, a strong debate occurred to determine whether DES used during "wide-awake surgery" with success is a gold standard to study brain functions (Mandonnet et al., 2010; Borchers et al. 2012; Desmurget et al., 2013). In this debate, confusion is very often made between DES and MES, as these are considered to induce similar effects on the nervous tissues, with comparable behavioural consequences. However, electrical stimulation (ES) parameters used in both techniques are clearly different. More surprisingly, a strong biophysical rational of their choices is lacking. It may be due to historical, methodological and technical constraints that have guided empirically them. These constraints may have strongly shaped and limited experimental protocols in a standard way. By contrast, the gap between MES and DES may reveal a great potential for new experimental paradigms in ES of the brain in vivo. By considering this gap and new technical developments in the design of stimulators, it may be time to move on to alternative and innovative stimulations protocols, especially regarding and inspired from what is performed in functional electrical stimulation (FES) of peripheral nerves, for which more theoretical supports exist.

5.1.7. Modelling of structural and functional brain connectivity networks in Diffuse Low-Grade Glioma (DLGG)

Participants: Jija Syamala James, Anirban Dutta, François Bonnetblanc, David Guiraud, Nicolas Menjot de Champfleur [Department of Neuroradiology,CHU, Montpellier], Emmanuelle Le Bars [Institute of Human Functional Imaging 12 FH, CHU, Montpellier], Hugues Duffau [Neurosurgery Department, CHU, Montpellier].

The impairment of functional brain connectivity networks in Diffuse Low Grade Glioma (DLGG) subjects can lead to distinct functional deficits where the challenge remains in greater understanding of distribution of dynamic brain connectivity. Our multimodality (resting state functional MRI, Diffusion Tensor Imaging and tractography) neuroimaging study aims to evaluate the differences in spatial and temporal patterns of brain connectivity networks in DLGG patients, pre and postoperatively.

In order to identify the brain connectivity networks, we analysed resting state fMRI data of 22 mesio-frontal DLGG patients by using MELODIC (Multivariate Exploratory Linear Optimised Decomposition into Independent Components)-ICA module, implemented in Freesurfer Software Library [FSL] (www.fmrib.ox.ac.uk/fsl). First we evaluated the clinical efficacy of rsfMRI technique to non-invasively map the dynamic functional reorganizations of brain connectivity networks in glioma subjects. Further we observed the existence of altered inter-hemispheric functional connectivity in DLGG subjects postoperatively.

We are currently performing this analysis using larger sample size in order to find out

a) the invivo structure-function relationship of motor, language and visual brain networks by fusing the information from DTI and fMRI.
b) the correlation between brain connectivity networks and neurobehavioural performance.
c) the neuroplastic alterations in topographic organization and strength of connections before and after DLGG surgery.

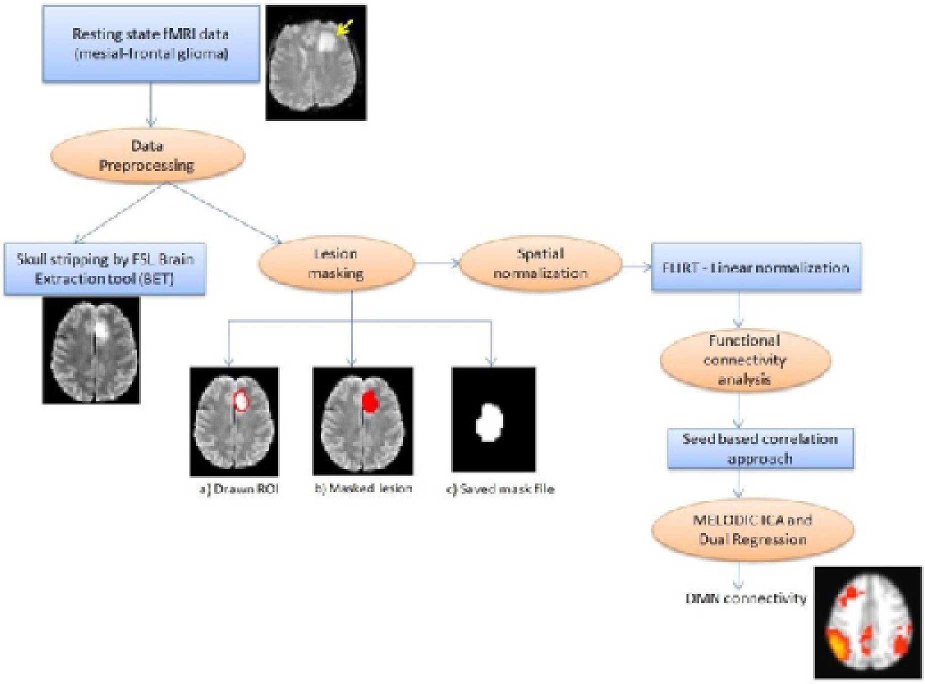


Figure 7. Steps involved in resting state functional connectivity analysis.

5.1.8. SPINSTIM: Direct spinal stimulation for rehabilitation of bladder, bowel and sexual functions in spinal cord injury

Participants: Christine Azevedo Coste, Luc Bauchet [Neurosurgery Department, CHU, Montpellier], Claire Delleci [CHU Bordeaux], Charles Fattal, Thomas Guiho, David Guiraud, Jean-Rodolphe Vignes [CHU Bordeaux].

For the general public, spinal cord injury (SCI) is often restricted to limb paralysis. In reality, by interrupting communication between encephalon and peripheral organs, medullary wounds lead to physiological deficiencies such as urinary (retention and/or incontinence), gastrointestinal (constipation) or sexual impairments; disorders which are the center of patient's expectations.

Spinal cord stimulation (SCS) is a general term which includes both epidural and intradural stimulation. Originally associated with the treatment of chronic neurogical pain (in the 1970ies), SCS led also to immediate and profound improvements of sensory and motor functions in recent studies both on SCI patients (few subjects involved) and rodents.

Despite these promising results some limitations have still to be overcome. Among them, the use of small animal models, the empirical aspect of the stimulation procedure and the impact of these protocols on intestinal and urinary functions are critical.

To counteract these limits, we want to explore intradural and epidural stimulations in an intermediate model- the house pig- and assess their impact on bladder, guts and genitals. In order to evaluate our approach, we will record EMG signals of lower limbs and sphincters (both urethral and anal), and simultaneously, we will monitor bladder and rectal pressure.

5.1.8.1. Development of point of care testing (POCT) device for neurovascular coupling from simultaneous recording of EEG and NIRS during non-invasive brain stimulation (NIBS) for closed-loop control of NIBS.

Participants: Mehak Sood, Utkarsh Jindal, Mitsuhiro Hayashibe, Stephane Perrey, Michael A. Nitsche, Abhijit Das, Shubhajit Roy Chowdhury, Anirban Dutta.

Our preliminary work showed that transcranial direct current stimulation (tDCS) can perturb local neuronal activity which can be used for assessing regional neurovascular coupling (NVC) functionality. It was postulated that tDCS leads to rapid dynamic variations of the brain cell microenvironment that perturbs the hemodynamic and electromagnetic responses. Based on these preliminary studies, we developed a POCT device for EEG-NIRS based screening and monitoring of neurovascular coupling functionality under perturbation with tDCS. The stroke case study showed detectable changes in the degree of NVC to a 0.526A/m^2 square-pulse (0-30sec) of anodal tDCS where these alterations in the vascular system may result in secondary changes in the cortical excitability. The objective of this case study was to evaluate an empirical method to assess NVC using cross-correlation function (CCF) between mean (cortical) tissue oxy-haemoglobin concentration time-series and averaged PSD time-course from the EEG spectrogram. The CCF based assessment of the patient-specific status of NVC are currently being studied in a larger cohort with small vessel diseases. The overarching goal is closed-loop control of tDCS based on simultaneous recording of EEG and NIRS during non-invasive brain stimulation.

5.2. Synthesis and Control of Human Functions

5.2.1. Ergonomics of the control by a quadriplegic of hand functions

Participants: Christine Azevedo Coste, David Guiraud, Wafa Tigra, Charles Fattal.

In subjects with complete Spinal Cord Injury (SCI) above C7, the four limbs are paralyzed (quadriplegia). Recovery of grasping movements is then reported as a priority. Indeed, most activities of daily living are achieved through upper limbs. Thus, restoration of hand and forearm active mobility could significantly increase independence and quality of life of these people and decrease their need of human aid. Although most of the subjects plebiscite pharmacological or biological solutions, only orthotics and Functional Electrical Stimulation (FES) allow, so far, to restore hand movements but they are rarely used. Ergonomics and comfort

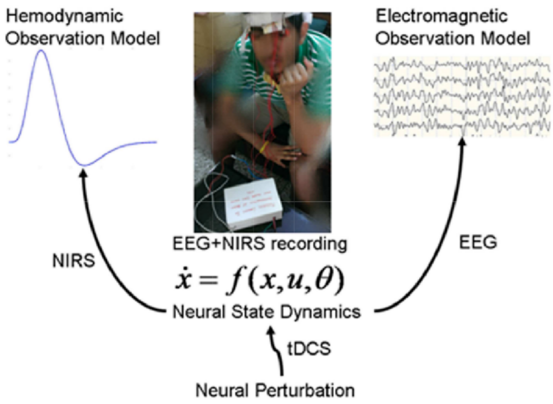


Figure 8. EEG-NIRS based simultaneous recording of the hemodynamic and electromagnetic responses to perturbations with transcranial direct current stimulation (tDCS).

of piloting mode could partly explain the low usage of these systems. In this context, our aim is to explore possible solutions for subjects to interact with such devices. In this article we propose to evaluate the capacity of active upper limb muscles contraction to be used to intuitively control FES in tetraplegic subjects. In this study, we assessed the ability to gradually contract different muscles: trapezius, deltoid, platysma and biceps. Three subjects with C6 to C7 neurological level of lesion were included. We show that over the active upper limb muscles tested, contraction of the trapezius muscle was considered by the subjects as the most comfortable and could be employed as an intuitive mode of control of functional assistive devices.

5.2.2. Implementation of filtering, calibration and reconstruction algorithms dedicated to the use of inertial measurement units related to rehabilitation and movement analysis

Participants: Christine Azevedo Coste, Benoît Sijobert, Roger Pissard-Gibollet.

This work has been done within SENSBIO ADT.

- Stride length estimation

  An algorithm has been implemented for the sep by step stride length calculation from only one shank located sensor algorithm (calibration, segmentation and reconstruction) Experimental validation was done on 10 healthy subjects (error <10%), 12 Parkinson disease subjects, and 7 hemiparetic subjects.
- Comparison of different reconstruction algorithms

  A Python programmed toolbox was developed for movement analysis (SensbioTK) in order to compare the performances of Mahony, Madgwick, Martin-Salaun sensor fusion algorithms.
- Real-time MoCap data processing Matlab based software

  The algorithm is based on Nexus SDK for rotation angle and translation calculation of a tracked object.

5.2.3. Dominant Component in Muscle Fatigue Induced Hand Tremor during Laparoscopic Surgical Manipulation

Participants: Sourav Chandra, Mitsuhiro Hayashibe, Asokan Thondiyath [IIT Madras].

Accuracy of laparoscopic surgery gets affected by the hand tremor of the surgeons. Though cognitive load is inevitable in such activity which promotes tremor, muscle fatigue induced tremor is significant among the most important source of tremor. Characteristic of fatigue induced hand tremor and its dominant directional properties are reported in this work. For a fixed laparoscopic tool grip with temporally synchronized predefined task protocols, characteristics of fatigue induced tremors have been studied. Dominant component of tremor is found to be in the sagittal plane in case of both static and dynamic tasks. The results shown in the figure, sagittal plane (z axis) component of hand tremor is higher than the other directions. In order to relate it with the muscle fatigue level, spectral properties of surface electromyography (SEMG) were also investigated simultaneously. A study of transient effect on tool positioning was also included, which conjointly advocates the other experimental results on fatigue induced hand tremor as well. Currently a better metric for muscle fatigue is being analyzed and studied with a purview of relation in between SEMG and hand tremor (fig.9).

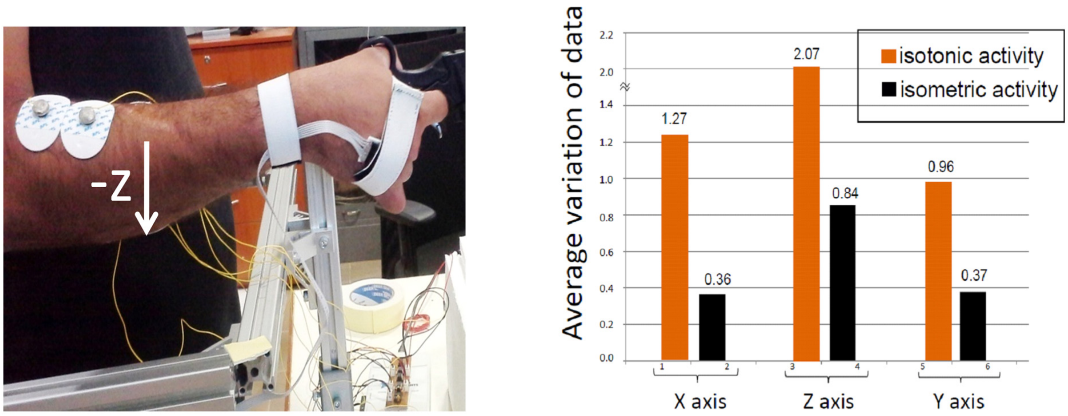


Figure 9. (left) Experimental hand tremor measurement setup (right) Dominant component of hand tremor in all three planes

5.2.4. Human arm optimal motion analysis in industrial task for improving ergonomics of painful workstations

Participants: Nahema Sylla, Vincent Bonnet [TUAT, Japan], Nahid Armande [PSA], Philippe Fraisse.

In PSA Peugeot Citroen factories, high precision requirements of workstations make them being manual. One of the main goal of the car manufacturer is to minimize the pain of workers while maintaining high efficiency of production lines. Consequently, assisting operators with an exoskeleton is a potential solution for improving ergonomics of painful workstations while respecting industrial constraints [33]. We have developped a new approach based on inverse optimization to better understand human arm motion in industrial screwing task. The process combines several criteria to minimize such as energy expenditure or trajectory smoothness leading to the optimal trajectory of a typical screwing task, often performed by workers. Estimated joint trajectories are similar with the measured ones, with a mean square error of 4 degrees. The resulting cost-function is mainly composed of energy expenditure and geodesic criteria. Results show the relevance of using composite cost function in human motion planning [52]. This method has been applied to evaluate a 7 DoF exoskeleton in terms of motion control. The results of our study show that the hybrid composition of the free arm movement was accurately determined. At contrary, when wearing the exoskeleton, which produces an arbitrary determined torque compensation, the motion is different from the naturally adopted one. This study is part of the evaluation and comprehension of the complex neuromuscular mechanism resulting in wearing an exoskeleton several hours per day for industrial tasks assistance [49].

5.2.5. A System for Real-time Online Estimation of Joint Torque with Evoked EMG under Electrical Stimulus

Participants: Zhan Li, Mitsuhiro Hayashibe, David Andreu, David Guiraud.

Functional electrical stimulation (FES) is a useful rehabilitation technique for restoring motor capability of spinal cord injured (SCI) patients by artificially driving muscle contraction through delivering electrical pulses. Real-time FES systems with online modulation ability are in great demand towards clinic applications. In this work, online estimation of joint torque with evoked electromyography (eEMG) in real-time environment is presented. The eEMG is acquired by National Instrument (NI) acquisition card and the stimulus is produced by wireless stimulator (manufactured by Vivaltis Inc., France). Kalman filter (KF) is adopted and embedded as the online estimator. Such real-time online torque estimation system produces promising results. Currently, the implementation to use EMG signal also from wireless Pod module is under investigation (fig.10).

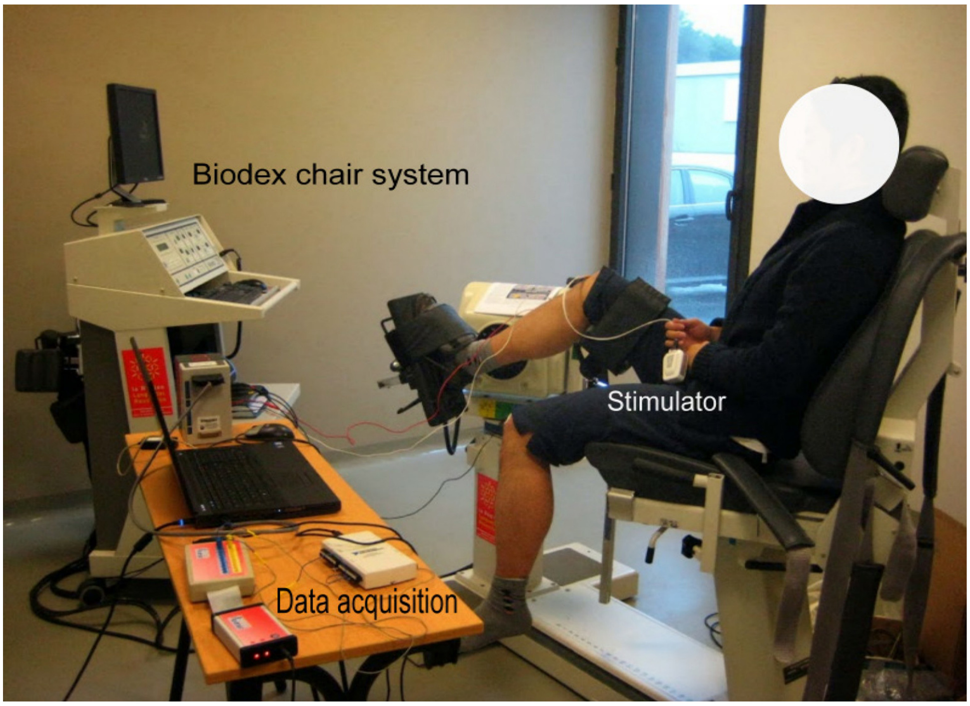


Figure 10. Real-time online FES-induced torque estimation system

5.2.6. Freezing of Gait Analysis and Detection

Participants: Christine Azevedo Coste, Benoît Sijobert, Roger Pissard-Gibollet, Christian Geny [CHU Montepllier, Neurologie].

We have extended and optimized the work on Freezing Of Gait (FOG) and destination detection (Maud Pasquier PhD thesis, 2013). A new software for FOG Criteria based on cadency and stride length calculation has been implemented. Results were compared to Freezing Index (Moore et al., 2008) based on frequencies analysis of legs vertical acceleration. The comparison between detection and reality is done on the basis of video analysis of the performed tasks. A software has been implemented for the video labelization Matlab/VLC based tool, with graphical user interface, for marking and synchronizing events from a video (MovieFOG). 14 Parkinson disease patients were included in the study [12][48] (fig.11).

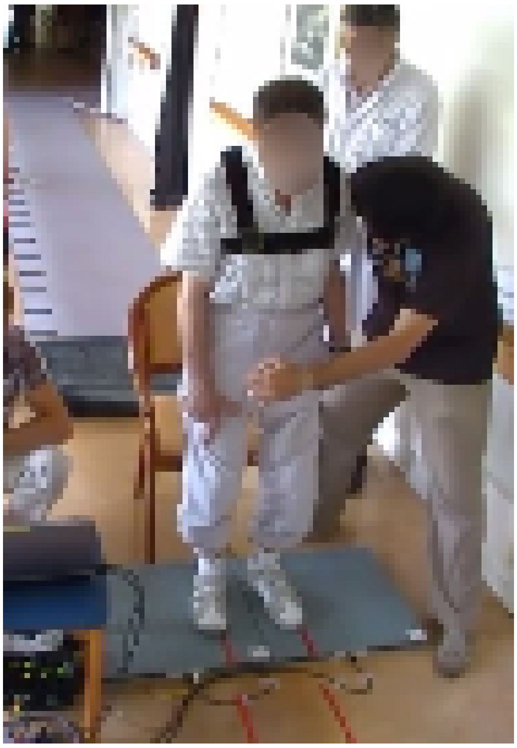


Figure 11. Freezing of Gait Analysis and Detection

5.2.7. A novel brain-computer interfacing paradigm for control of multi-DOFs robot with adaptive EEG decoding and synergetic environment adaptation

Participants: Saugat Bhattacharyya, Shingo Shimoda [RIKEN, Japan], Mitsuhiro Hayashibe.

The study proposes a novel brain-computer interfacing paradigm for control of a multi-joint redundant robot system. Here, the user would determine the direction of end-point movement of a 3-dof robot arm using motor imagery electroencephalography (EEG) signal with co-adaptive decoder while a synergetic motor learning algorithm manages a peripheral redundancy in multi DOF joints toward energy optimality through tacit learning. As in human motor control, torque control paradigm is employed for a robot to be sensitive to the given environment. The dynamic condition of the robot arm is taken into consideration by the learning algorithm. Thus, the user needs to only think about the end-point movement of the robot arm, which allows simultaneous multi-joints control by BCI. The k-Nearest Neighbor based decoder designed for this study is adaptive to the changing mental state of the user. Offline experiments on the decoder reveals that its classification accuracy gradually increases at each learning stage. Online experiments also reveals that the users successfully reach their targets with an average decoder accuracy of more than 65% in different end-point load conditions. The details of the BCI control paradigm, shown in Fig.12 , is as follows: Initially, the robot is trained to its dynamic environment using a tacit learning approach for a fixed period of time. In this study, the load carried by the robot is treated as the environmental changes along with link segment inertial configuration changes. As a result, the movement of the joints of the robot adapts to the changing load. After the training of the robot, the subject begins his/her task of visualizing the target and decides on the direction of motion of the robot. Here, we have used left and right movement imagery to move the robot in upward and downward direction, respectively. Subsequently, the EEG signals (corresponding to the movement imagined) are pre-processed and their corresponding features are extracted. These features are used as inputs to the decoder which determines the mental state of the user and sends command to the robot to move in the equivalent direction.

5.2.8. Impact of the gaze direction on the skier trajectory

Participants: Christine Azevedo Coste, Benoît Sijobert, Roger Pissard-Gibollet, Nicolas Coulmy [FFS annecy].

This work is done within a collaboration with the French Skiing Federation (FFS). Preliminary indoor tests for designing an experimental protocol using motion capture and inertial sensors was realized using a ski simulator and Motion Capture Tool from DEMAR (VICON BONITA). Indoor and outdoor experimentations were done with olympic skiers. In outdoor tests IMUs were combined with SMI eye tracking device. Preliminary analysis and data reconstruction has been done. A VPython script for 3D visualizing of skier movements was developed (fig.13).

5.2.9. Development of a low-cost biofeedback system for electromyogram-triggered functional electrical stimulation therapy in conjunction with non-invasive brain stimulation

Participants: Anirban Dutta, Christine Azevedo Coste, Mitsuhiro Hayashibe, Uttama Lahiri, Abhijit Das, Michael A. Nitsche, David Guiraud.

Functional electrical stimulation (FES) facilitates ambulatory function after paralysis by activating the muscles of the lower extremities. The FES-assisted stepping can either be triggered by a heel-switch, or by an electromyogram-(EMG-) based gait event detector. A group of six chronic (>6 months post-stroke) hemiplegic stroke survivors underwent transcutaneous FES-assisted training for 1 hour on stepping task with EMG biofeedback from paretic tibialis anterior (TA) and medial gastrocnemius (GM) muscles, where the stimulation of the paretic TA or GM was triggered with surface EMG from the same muscle. During the baseline, post-intervention, and 2-day-postintervention assessments, a total of 5 minutes of surface EMG was recorded from paretic GM and TA muscles during volitional treadmill walking. Two-way ANOVA showed significant effects in terms of P-values for the 6 stroke subjects, 0.002, the 3 assessments, 0, and the interaction between subjects and assessments, 6.21E-19. The study showed a significant improvement from baseline in paretic GM and TA muscles coordination during volitional treadmill walking. Moreover, it was found that the EMG-triggered

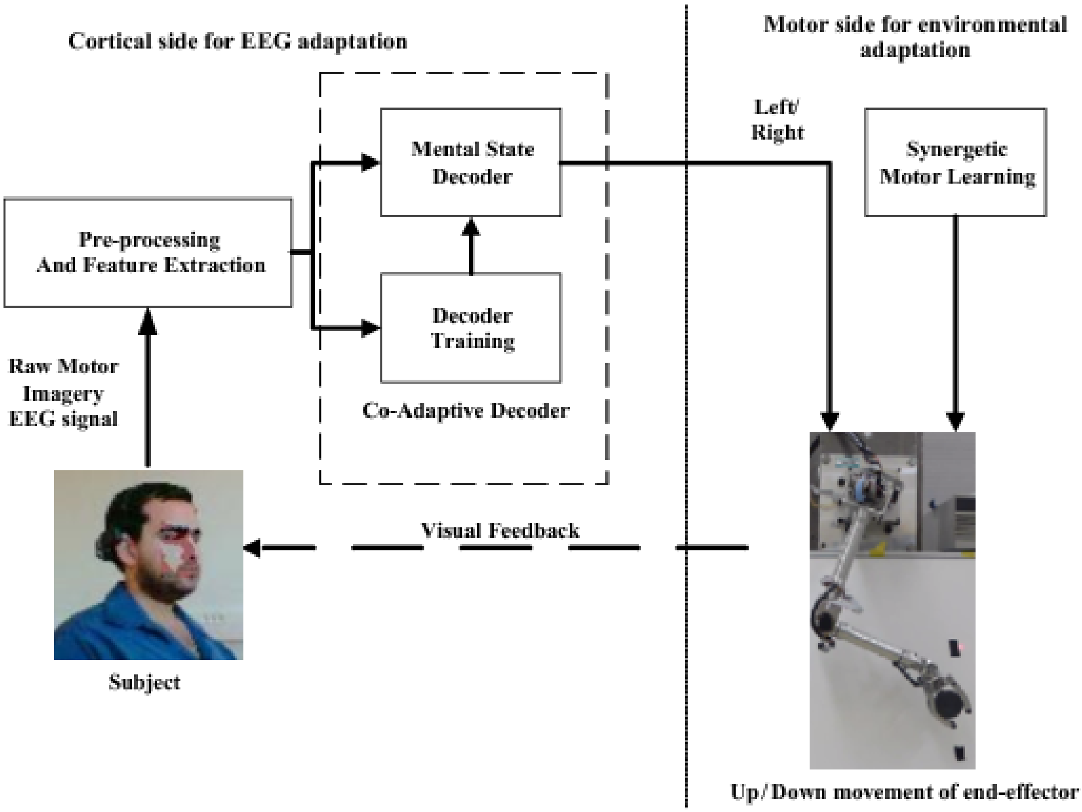


Figure 12. BCI paradigm employed in this study for control of multi-DOFs robot using adaptive left-right motor imagery decoder and synergetic motor learning for peripheric redundancy management (via tacit learning).

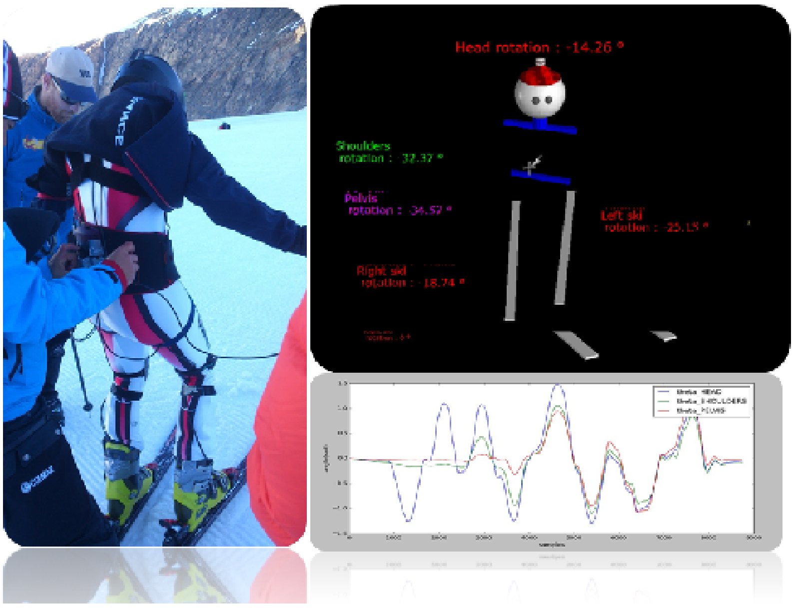


Figure 13. Impact of the gaze direction on the skier trajectory

FES-assisted therapy for stand-to-walk transition helped in convergence of the deviation in centroidal angular momentum from the normative value to a quasi-steady state during the double-support phase of the nonparetic. Also, the observational gait analysis showed improvement in ankle plantarflexion during late stance, knee flexion, and ground clearance of the foot during swing phase of the gait. Currently, we are conducting preliminary stroke studies to evaluate non-invasive brain stimulation as an adjunct to EMG-triggered FES therapy for movement rehabilitation [19][22].

5.2.10. Development and bench-testing of a low-cost eye tracking system (ETS) to measure gaze abnormality in stroke towards virtual reality based visuomotor therapy task.

Participants: Deepesh Kumar, Abhijit Das, David Guiraud, Michael A. Nitsche, Anirban Dutta, Uttama Lahiri.

We conducted a preliminary usability study while incorporating our novel low-cost ETS to measure one's eye gaze indices in response to presented visual task. The ETS provided gaze-related biomarkers which has the potential to be mapped to the probable abnormalities in one's eye movement pattern in stroke. Our preliminary findings with stroke-survivors and age-matched healthy participants indicate the potential of our low-cost ETS to provide quantitative measures of the difference in gaze-related biomarkers between the two groups of participants. Based on these preliminary results, we are conducting a clinical stroke study on ETS based screening and monitoring of performance during a virtual reality based visuomotor balance therapy task.

5.3. Neuroprostheses and technology

5.3.1. Abstraction and composition for formal design of neuroprotheses

Participants: David Andreu, Hélène Leroux, Karen Godary-Dejean.

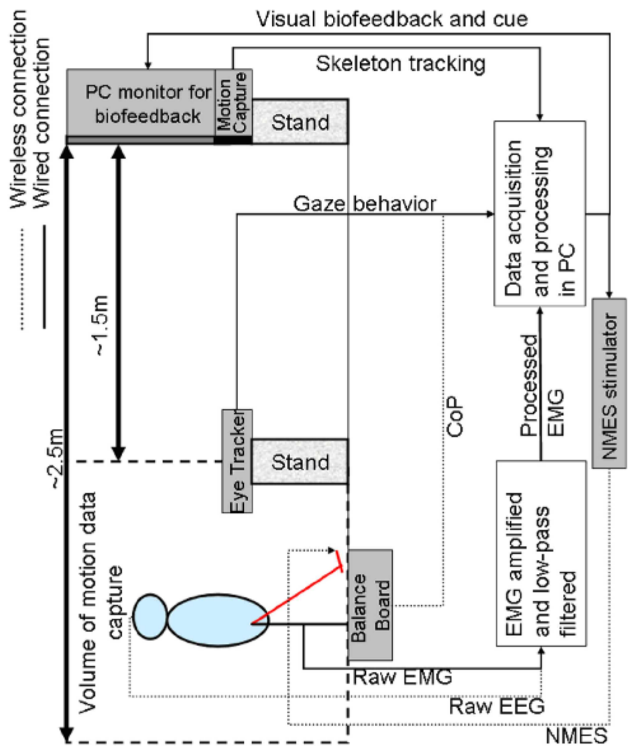


Figure 14. Human-machine-interface integrating low-cost sensors for post-stroke balance rehabilitation

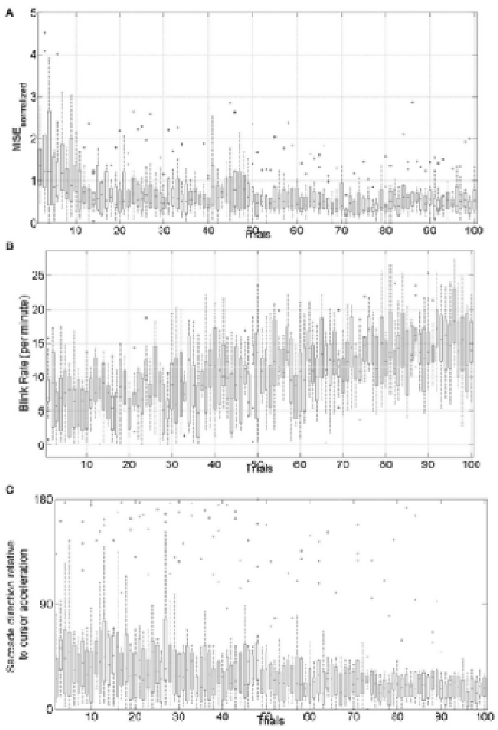


Figure 15. A. Box-plot of normalized mean squared error (MSE) across 10 healthy subjects, B. Box-plot of the blink rate during the visuomotor task, C. Box-plot of saccadic direction relative to the cursor acceleration during the visuomotor task.

In the framework of specification and implementation of complex digital systems on FPGA, we have developed an approach based on components whose behavior and composition are specified by generalized interpreted T-time Petri nets (HILECOP). One of the inherent difficulties for designer is, on the behavioral part, to account for exceptions. This often leads to a complex modeling and is a source of human errors. Indeed, it is intricate to express all the possible situations (i.e. current state of model). We have defined a way to model exception handling by integrating the well-know concept of macroplace into the formalism. The analysability of the model and the efficiency of the implementation on FPGA (reactivity and surface, ie number of logic blocks) have been preserved. An example of macroplace is given in figure 13; it contains a sub-net (set of places of its refinement) from which exception handling is simply described by a dedicated output transition (transition te on fig. 13), whatever is the current state of the sub-net.

We also solved state evolution conflicts introducing (automatically) priorities between transitions, to avoid reaching inconsistent global state while synchronously executing the model.

The new formalism (including all improvements) has been defined [45], as well as the model transformation based equivalent PNML generation for using existing analysis tools. The VHDL code generation has also been defined [59].

All this work has been applied to an industrial example, that of a neural stimulator developed in collaboration with MXM industrial partner. Results have shown the significant contribution of the theoretical approach to the stimulation device reliability, while preserving both surface and power consumption of the given digital part of the device.

Ongoing work, developed through I. Merzoug PhD thesis, concerns the improvement of the analysis of synchronously implemented Petri nets.

5.3.2. New FES dedicated digital processor for neurostimulator

Participants: David Andreu, David Guiraud.

We designed (patent pending) and prototyped a new neural FES dedicated processor and its associated (more compact and efficient) set of instructions, as well as an embedded sequencer for accurate timing in sequencing stimulations to be performed (by the stimulator). The new neural stimulator is based on a dedicated ASIC (Application Specific Integrated Circuit), that is able to drive up to 24 channels of stimulation in absolute synchronization, and with a programmable and controlled current level distribution (patent pending). This ASIC also allows for impedance measurement. The functions of the stimulator are currently implemented in two separate chips: an analog stimulation front-end (ASIC) and a field-programmable gate array (FPGA) embedding the logic control. The FPGA embeds the new FES dedicated processor setting the output stage configuration (poles configuration and current ratio between them) and running potentially complex stimulation profiles (with a 1μs time step and 5 μA current step); example of generated stimulations are shown in Fig.16 . It also embeds the protocol stack allowing for remote programming and online control. Online control relies on advanced and efficient modulation mechanisms, e.g. coefficient based modulation preserving balanced stimulation.

And last but not least, it also embeds a monitoring module ensuring the respect of safety constraints stemming both from target tissue protection and electrode integrity preservation; this reference model based monitoring module ensures (configurable) current and quantity of injected charges limits and thus safe stimulation whatever are electrodes to be used (particularly for thin-film micro-electrodes). Safety limits must be defined by users (partners) according to the target and electrodes to be used.

5.3.3. Fast simulation of hybrid dynamical systems

Participants: Abir Ben Khaled [IFPEN], Daniel Simon, Mongi Ben Gaid [IFPEN].

When dealing with the design of complex systems, simulation is an indisputable step between concept design and prototype validation. Realistic simulations allow for the preliminary evaluation, tuning and possibly redesign of proposed solutions ahead of implementation, thus lowering the risks. To be confident in the result, building such simulations needs high fidelity models both for the components and for their interaction. Models

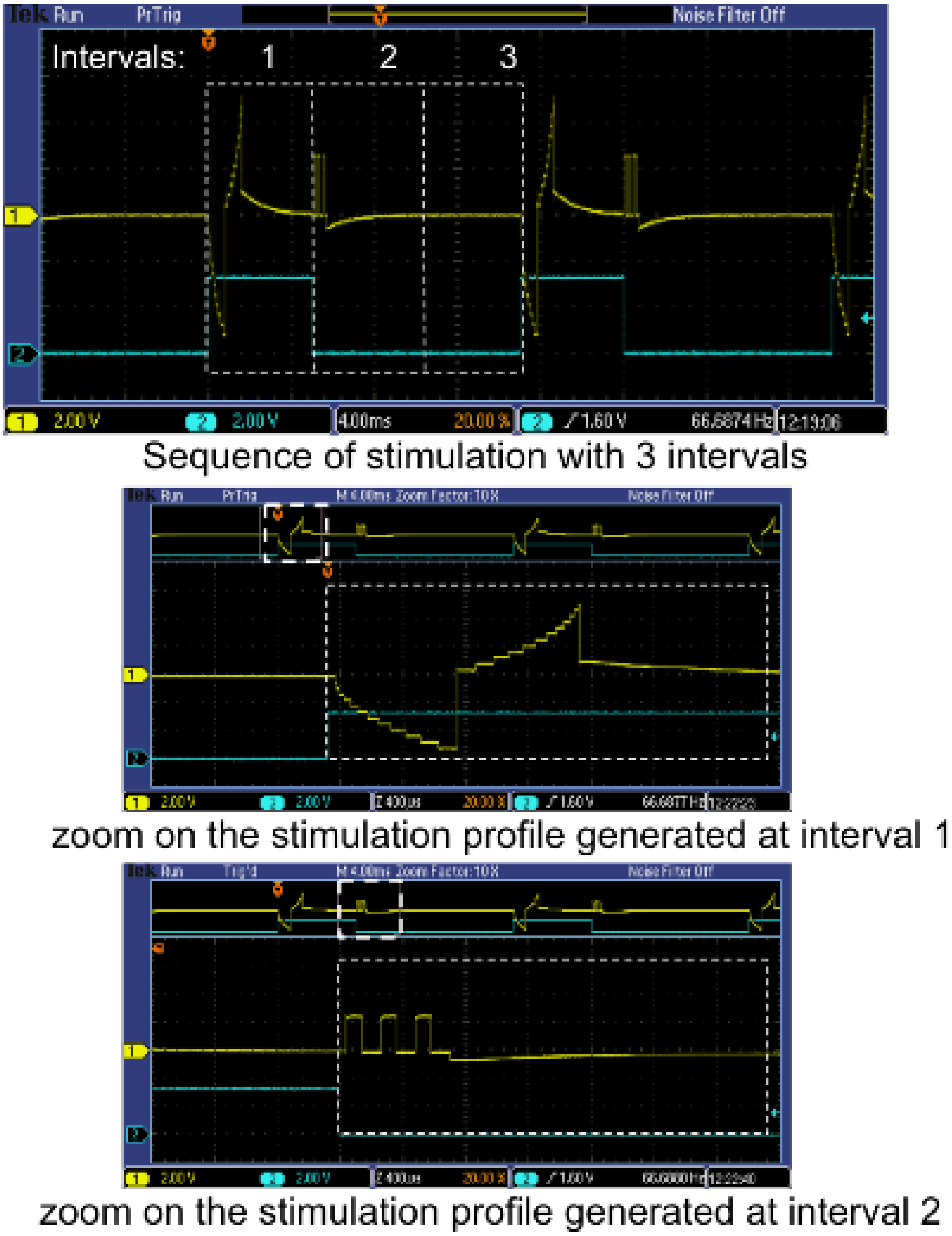


Figure 16. Sequence of stimulation generated by the new neurostimulator

of dynamical systems (as, for example, muscular fibers) are often given as a set of Ordinary Differential Equations (ODEs). However, the simulation of high-fidelity models is time consuming, and reaching real-time constraints is out of the capabilities of monotithic simulations running on single cores.

The aim of the on-going work is to speed up the numerical integration of hybrid dynamical systems, eventually until reaching a real-time execution, while keeping the integration errors inside controlled bounds. The basic approach consists in splitting the system into sub-models, which are integrated in parallel. I has been shown that an efficient partition must minimize the interactions between sub-models, in particular by confining discontinuities processing inside each component. Automatic partitioning, based on some particular incidence matrices of the original system, has been investigated to introduce a finely-grained co-simulation method enabling numerical integration speed-ups [14]. It is obtained using a partition across the model into loosely coupled sub-systems with sparse communication between modules. The proposed scheme leads to schedule a large number of operations with a wide range of execution times. A suitable off-line scheduling algorithm, based on the input/output dynamics of the models, is proposed to minimize the simulation errors induced by the parallel execution.The method was tested with an automotive engine model, but it is generic and can be applied to other systems of hybrid ODEs/DAEs, as are large sets of muscular fibers.

However, slack synchronization intervals may generate integration errors in the final result. Rather than using costly small integration and communication steps, an enhanced method uses context-based extrapolation is investigated to improve the trade-off between integration speed-ups, needing large communication steps, and simulation precision, needing frequent updates for the models inputs. the method uses extrapolations of the behavior of the models over the synchronization intervals. Test results on a hybrid dynamical engine model, based on FMI for model exchange, show that well chosen context-based extrapolation allows for significant speed-up of the simulation with negligible computing overheads [37].

5.3.4. Control and scheduling co-design for stimulation systems

Participants: Daniel Simon, David Andreu, Samy Lafnoune [Master2 Robotique].

In the FES distributed system developed by the team and marketed by Vivaltis, external electrodes are sticked on the body to either stimulate a muscle or to measure sensitive information (e.g., EMG). Each electrode is connected to a pod, which can be either a stimulation pod or a measurement pod. These Pods are controlled and coordinated by a controler through wireless radio-frequency (RF) links.

The communication frames between the controler and the pods use frames, each frame contains the receiver address and a code which correspond to the deired action. Simple communication stacks, based on the reduced OSI model, are implemented in each node of the wireless network, and the current version only provides static scheduling of the messages.

In fact static scheduling is not at all optimal when the generation of complex motions needs a tight coordination between several sensors and actuators over closed-loop controllers. The high sensitivity of the RF link with respect to varying networking loads and environment conditions calls for an adaptive scheduling of the messages via the regulation of a QoS criterion. The design of such feedback controllers wil rely on previous work [53], [54].

As any modification of the existing devices can be costly and as experiments involving livings cannot be done easily, a real-time simulation system has been design and implemented. The system includes continuous models of the muscular and skeleton systems, models of the wireless network, simplified communicaion stacks and control code running inside real-time threads. The system is open, running C code inside posix threads under Linux, so that the models can be progressively detailed and enriched when necessary [56].

DIANA Team

5. New Results

5.1. Highlights of the Year

Arnaud Legout and Thierry Parmentelat designed and realized the very first Inria Mooc hosted on the FUN platform. This Mooc is devoted to the study of the Python language, and targets undergrandudate students. The objective of the course is to give students a thorough understanding of the internal mechanisms of language, and lead them to small and realistic applications. This Mooc was a big success: 9166 persons registered to the course, out of them five hundred followed the whole course and more than a hundred finished the project. For more details on this Mooc see https://www.france-universite-numerique-mooc.fr/courses/inria/41001/Trimestre_4_2014/about.

5.2. From network-level measurements to expected QoE: the Skype use case

Contributors: Salim Afra, Chadi Barakat and Damien Saucez. Applications rely on rich multimedia contents and experience of end users is sensitive to network conditions. Consequently, network operators must design their infrastructure to ensure high Quality of Experience (QoE) for their customers. However, applications are usually over-the-top services on which network operators have no control and users have no mean to tune the network when they undergo poor QoE. In this project, called ACQUA for Application for the Prediction of Quality of Experience at Internet Access, we propose a new approach that allows network operators to determine how their network performance will influence QoE and end users to predict the QoE even before launching their applications. We predict the subjective QoE users will undergo based on the knowledge of objective network performance parameters obtained with active measurements (e.g., delay, loss) and machine learning. With the particular case of Skype calls and using a decision tree, we show that our approach achieves 83% of accuracy when estimating QoE from the delay, bandwidth, and loss. Our approach can be seen as a new way of performing measurements at the Internet access, where instead of expressing the expected performance in terms of network-level measurements, the performance of the access is expressed in clear terms related to the expected quality for the main applications of interest to the end user. The strength of the approach is in its capacity of expressing directly the QoE as a function of network-level measurements, which is an enabler for QoE prediction, and in reusing the same network-level measurements as input to different models for the QoE of end user applications. More details on this approach and on our application ACQUA can be found in section 4.5 , in the report summarizing the results [24] and on the application web page http://team.inria.fr/diana/acqua/.

5.3. Understanding of modern web traffic

Contributors: Salim Afra, Chadi Barakat, Byungchul Park and Damien Saucez.

Mobile devices are everywhere nowadays but little is known about the way they differ from traditional non-mobile devices in terms of usage and the characteristics of the web traffic they generate. In this contribution, we propose a first study of the differences that exist between mobile and non-mobile Web traffic seen from the lorgnette of a university campus network. The study is performed at different levels starting from users' behavior to transport protocol configurations. Our main findings are that mobile users often browse websites tailored to their devices. They show a significant adoption of Apps to browse the web and a preference for multimedia content. The different way of conceiving the web for mobiles is reflected at the HTTP and TCP levels with much less HTTP redirections and abrupt TCP connection terminations. Interestingly, mobile traffic carries larger contents and have larger TCP flows than non-mobile traffic. By cross-analysis of protocols and users' behavior, we explain why TCP flows in mobile traffic are larger than those of non-mobiles. Further details on this study can be found in [30].

5.4. Characterizing ICMP Rate Limitation on Routers

Contributors: Chadi Barakat and Ricardo Ravaioli.

In the last decade, path discovery has been extensively covered in the literature. In its simplest form, it generally works by sending probes that expire along the path from a host to a destination. It is also known that network administrators often configure their routers to limit the amount of ICMP replies sent, a common practice typically referred to as ICMP rate limitation. In this contribution we attempt to characterize the responsiveness of routers to expiring ICMP echo-request packets. Our contribution is twofold: first, we provide a detailed analysis of how routers are most commonly configured to respond to expiring packets; next, we show that for the vast majority of routers the measured round-trip time is not affected by the probing rate. This contribution is published in ICC'2015 [21]. It is the result of a collaboration with the SIGNET group at I3S in the context of a PhD thesis funded by the UCN@SOPHIA Labex.

5.5. Studying Social Networks at Scale: Macroscopic Anatomy of the Twitter Social Graph

Contributors: Maksym Gabielkov and Arnaud Legout.

Twitter is one of the largest social networks using exclusively directed links among accounts. This makes the Twitter social graph much closer to the social graph supporting real life communications than, for instance, Facebook. Therefore, understanding the structure of the Twitter social graph is interesting not only for computer scientists, but also for researchers in other fields, such as sociologists. However, little is known about how the information propagation in Twitter is constrained by its inner structure. We have performed an in-depth study of the macroscopic structure of the Twitter social graph unveiling the highways on which tweets propagate, the specific user activity associated with each component of this macroscopic structure, and the evolution of this macroscopic structure with time for the past 6 years. For this study, we crawled Twitter to retrieve all accounts and all social relationships (follow links) among accounts; the crawl completed in July 2012 with 505 million accounts interconnected by 23 billion links. Then, we proposed a methodology to unveil the macroscopic structure of the Twitter social graph. This macroscopic structure consists of 8 components defined by their connectivity characteristics. Each component group users with a specific usage of Twitter. For instance, we identified components gathering together spammers, or celebrities. Finally, we presented a method to approximate the macroscopic structure of the Twitter social graph in the past, validate this method using old datasets, and discuss the evolution of the macroscopic structure of the Twitter social graph during the past 6 years. This study was published in ACM Sigmetrics 2014 [17].

5.6. When AIMD meets ICN: a bandwidth sharing perspective

Contributors: Chadi Barakat and Damien Saucez.

Information-centric networking (ICN) leverages content demand redundancy and proposes in-network caching to reduce network and servers load and to improve quality of experience. In this contribution, we study the interaction between in-network caching of ICN and Additive Increase Multiplicative Decrease (AIMD) end-to-end congestion control with a focus on how bandwidth is shared, as a function of content popularity and cache provisioning. As caching shortens AIMD feedback loop, the download rate of AIMD is impacted. Supported by an analytical model based on Discriminatory Processor Sharing and real experiments, we observe that popular contents benefit from caching and realize a shorter download time at the expense of unpopular contents, which see their download time inflated by a factor bounded by $1/(1-\rho)$, where ρ is the network load. This bias can be removed by redefining congestion control to be delay independent or by over-provisioning link capacity at the edge so that to compensate for the greediness of popular contents. Further details on this study, which is the result of a collaboration with Politecnico di Bari, can be found in [23].

5.7. On the incentives and incremental deployments of ICN technologies for OTT services

Contributors: Chadi Barakat and Damien Saucez.

With the explosion of broadband Over-The-Top (OTT) services, the Internet is autonomously migrating toward overlay and incrementally deployable content distribution infrastructures. Information-Centric Networking (ICN) technologies are the natural candidates to efficiently distribute popular content to users. However, the strategic incentives in exploiting ICN, for both users and ISPs, are much less understood to date. We hence studied in [15] the strategic incentives for ICN overlay adoption in OTT services based on a game theoretical approach and discussed how OTTs shall shape their prices to motivate ICN overlay usages.

5.8. On ICN Cache Allocation to Content Providers

Contributor: Damien Saucez
Cross-Team Contributors: Mahmoud El Chamie (Maestro)
External contributors: Sahar Hoteit and Stefano Secci from Sorbonne Universités, UPMC Univ Paris 06.

Information Centric Networks (ICNs) allow offloading content distribution from content service providers by means of in-network caching. Despite a rather high maturation in the definition of ICN forwarding techniques, minor attention has been given to the strategic interaction among the multiple ICN stakeholders. We decided to focus on situations involving multiple Content Providers (CPs) and one ICN provider having to give them access to its caches. Intuitively, this situation is prone to high cache contention, in particular at the appealing topology cross-points. To address this problem we propose a resource allocation and pricing framework to support the network provider in the cache allocation to multiple CPs, for situations where CPs have non-overlapping sets of files and untruthful demands need to be avoided. As cache imputations to CPs need to be fair and robust against overclaiming, we evaluated common proportional and max-min fairness (PF, MMF) allocation rules, as well as coalitional game rules, the Nucleolus and the Shapley value. We found that the naive least-recently-used-based ICN approach provides proportional fairness. Moreover, the game-theoretic rules outperform in terms of content access latency the naive ICN approach as well as PF and MMF approaches, while sitting in between PF and MMF in terms of fairness. This paper is under submission [27].

5.9. Demonstrating a unified ICN development and evaluation framework

Contributors: Walid Dabbous, Alina Quereilhac, Damien Saucez and Thierry Turletti.

Information-Centric Networking solutions target world-wide deployment in the Internet. It is hence necessary to have access to a development and evaluation environment which enables both controllable and realistic experimentation to thoroughly understand how ICN solutions would behave in real life deployment. Such solution can be obtained with NEPI that we demonstrated at the ACM Information Centric Networking 2014 conference. In this demonstration, we presented a development and evaluation framework that combines emulation and live prototyping environments to provide ICN designers and implementers the means to build beyond-prototype ICN solutions. This framework is built upon NEPI. We demonstrated the benefits of such integrated approach by showing how complete experimental studies can be carried out with minimum manual intervention and experiment set-up overhead, in both emulation and live environments. More precisely, we demonstrated how to deploy the same experiment in different environment and how NEPI can help to minimise the implementation and operational overhead. This demonstration is summarised in [31].

5.10. Optimizing rules placement in OpenFlow networks: trading routing for better efficiency

Contributors: Chadi Barakat, Xuan Nam Nguyen, Damien Saucez and Thierry Turletti

The idea behind Software Defined Networking (SDN) is to conceive the network as one programmable entity rather than a set of devices to manually configure, and OpenFlow meets this objective. In OpenFlow, a centralized programmable controller installs forwarding rules onto switches to implement policies. However, this flexibility comes at the expense of extra overhead as the number of rules might exceed the memory capacity of switches, which raises the question of how to place most profitable rules on board. Solutions proposed so far strictly impose paths to be followed inside the network. We advocate instead that we can relax routing requirements within the network to concentrate on the final destination to which the traffic should forwarded, not how to route to this destination. In [19] we illustrate the concept, with an optimization problem that gets the maximum amount of traffic delivered according to policies and the actual dimensioning of the network. The traffic that cannot be accommodated is forwarded to the controller that has the capacity to process it further. [19] also demonstrates that our approach permits a better utilization of scarce resources in the network. We extended the work by stating that in many situations (e.g., data-center networks), the exact path followed by packets has not significant impact on performances as long as packets are delivered to their final destination decided by the endpoint policy. It is thus possible to deviate part of the traffic to alternative paths so as to better use network resources without violating the endpoint policy. In [20], we propose a linear optimization model of the rule allocation problem in resource constrained OpenFlow networks with loose routing policies. We show that the general problem is NP-hard and propose a polynomial time heuristic, called OFFICER, that aims at maximizing the amount of carried traffic in under-provisioned networks. Our numerical evaluation on four different topologies show that exploiting various paths allows to increase the amount of traffic supported by the network without significantly increasing the path length.

5.11. A Survey of Software-Defined Networking

Contributors: Bruno Astuto Arouche Nunes, Xuan Nam Nguyen and Thierry Turletti.

We wrote a survey of the emerging field of Software-Defined Networking (SDN). SDN is currently attracting significant attention from both academia and industry. Its field is quite recent, yet growing at a very fast pace. Still, there are important research challenges to be addressed. We look at the history of programmable networks, from early ideas until recent developments. In particular we described the SDN architecture in detail as well as the OpenFlow standard. We provided an overview of current SDN implementations and testing platforms and examined network services and applications that have been developed based on the SDN paradigm. We concluded with a discussion of future directions enabled by SDN ranging from support for heterogeneous networks to Information Centric Networking (ICN). The survey has been published in the IEEE Surveys and Tutorials journal [9]. This paper is among the top downloads on IEEE Exlpore in December 2014. See http://ieeexplore.ieee.org/xpl/browsePopular.jsp?reload=true.

5.12. Software-Defined Networking Enabled Capacity Sharing in User Centric Networks

Contributors: Bruno Astuto Arouche Nunes and Thierry Turletti.

We proposed to use SDN to deploy capacity sharing mechanisms in the context of User Centric Networking (UCN). We consider user-centric networks as a way of considerably mitigating the problem of sharing limited network capacity and resources efficiently and in a fairly manner. UCNs are self-organizing networks where the end user plays an active role in delivering networking functions such as providing Internet access to other users. We propose to leverage the SDN paradigm to enable cooperation between wireless nodes and to provide capacity sharing services in UCNs. Our proposed approach allows coverage of existing network infrastructure (e.g., WiFi or 3GPP) to be extended to other end users or ad hoc networks that would otherwise not be able to have access to network connectivity and services. Moreover, it takes into account current network load and conditions, and QoS requirements of applications. This work has been published in a special issue of Communications Magazine [14].

5.13. Decentralizing SDN's Control Plane

Contributors: Bruno Nunes Astuto and Thierry Turletti.

Motivated by the internets of the future that will likely be considerably larger in size as well as highly heterogeneous and decentralized, we sketched out a framework aiming to enable not only physical, but also logical distribution of the Software-Defined Networking (SDN) control plane. This framework will accomplish network control distribution by defining a hierarchy of controllers that can “match” an internet’s organizational– and administrative structure. The main idea is to delegate control between main controllers and secondary controllers in order to accommodate administrative decentralization and autonomy. This work has been presented in a short paper at the IEEE LCN conference [22].

5.14. Extending DCE to emulate Wireless Software Defined Networks.

Participants: Emilio Mancini, Hardik Soni, Thierry Turletti and Walid Dabbous.

Today it is not possible to simulate and evaluate in a realistic way wireless Software Defined Networking solutions. Indeed, the most used SDN emulator tool, Mininet, can only emulate point-to-point physical links using virtual Ethernet pairs (e.g., MAC layer is ignored), and it cannot provide mobility models for wireless nodes. To make the Direct Code Execution module (DCE) able to run Software Defined Networks we started to support OpenFlow NOX controller and Open vSwitch. The actual NOX binary is executed on a simulated ns-3 node. OpenFlow wireless routers are simulated using the Open vSwitch distribution with data-path kernel module support as it is widely used. DCE provides a mechanism to incorporate such a kernel module based application execution. A demonstration has been done at the 17th ACM International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems [29].

5.15. On the Performance of the LISP Beta Network

Contributor: Damien Saucez

The future Internet has been a hot topic during the past decade and many approaches towards this future Internet, ranging from incremental evolution to complete clean slate ones, have been proposed. One of the proposition, LISP, advocates for the separation of the identifier and the locator roles of IP addresses to reduce BGP churn and BGP table size. Up to now, however, most studies concerning LISP have been theoretical and, in fact, little is known about the actual LISP deployment performance. In [16], we report the measurement campaigns carried out on the LISP Beta Network. More precisely, we evaluated the performance of the two key components of the infrastructure: the control plane (i.e., the mapping system) and the interworking mechanism (i.e., communication between LISP and non-LISP sites). Our measurements highlight that performance offered by the LISP interworking infrastructure is strongly dependent on BGP routing policies. If we exclude misconfigured nodes, the mapping system typically provides reliable performance and relatively low median mapping resolution delays. Although the bias is not very important, control plane performance favors USA sites as a result of its larger LISP user base but also because European infrastructure appears to be less reliable.

This work resulted in a collaboration with Telecom ParisTech starting in mid-2014 a PhD thesis on the feasibility of large scale measurement of LISP networks with Luigi Iannone as advisor and Damien Saucez as co-advisor.

5.16. Standardization: Contributions to the IETF LISP WG

Contibutor: Damien Saucez

In the context of the LISP WG, we contributed to an Internet-draft called "An Architectural Introduction to the LISP Location-Identity Separation System" [25] that describes the architecture of the Locator/ID Separation Protocol (LISP), making it easier to read the rest of the LISP specifications and providing a basis for discussion about the details of the LISP protocols. This document is used for introductory purposes, more details can be found in RFC6830, the protocol specification. This internet-draft is in RFC queue, for imminent publication as RFC.

In the context of the LISP WG, we contributed to an Internet-draft called "LISP Threats Analysis" [33] that proposes a threat analysis of the Locator/Identifier Separation Protocol (LISP). This internet-draft is under discussion in the Working Group.

In the context of the LISP WG, we contributed to an Internet-draft called "LISP-Security (LISP-SEC)" [28] that specifies LISP-SEC, a set of security mechanisms that provides origin authentication, integrity and anti-replay protection to LISP's EID-to-RLOC mapping data conveyed via mapping lookup process. LISP-SEC also enables verification of authorization on EID-prefix claims in Map-Reply messages. This internet-draft is under discussion in the Working Group.

In the context of the LISP WG, we contributed to an Internet-draft called "LISP Impact" [32]. The Locator/Identifier Separation Protocol (LISP) aims at improving the Internet scalability properties leveraging on three simple principles: address role separation, encapsulation, and mapping. In this internet-draft, based on implementation, deployment, and theoretical studies, we discuss the impact that deployment of LISP can have on both the Internet in general and for the end-users in particular. This internet-draft is adopted as Working Group document on December 2014.

DICE Team

5. New Results

5.1. The economy of intermediation

We have presented in [6] an introductory panorama on the disruption of the intermediation revolution. Our efforts to measure data flows in the world, have been pursued [2] to estimate the concentration of the data industry. It is well known that the main platforms of the Web are concentrated in a few countries, mostly in the USA. Some countries, mostly in Asia, such as China, Russia, Korea or Japan have successfully developed their own Web 2.0 industry, while others, such as European countries, have failed to do so. We have explored in [7] the strategy of China, which has the largest Web industry behind the US and has made a priority of keeping its data at home, with systems in all activity sectors developed in general only one or two years after their main American counterparts. The innovation strategy of China aims in all fields to achieve technological independence, with at most 30% of foreign IP.

The rise of the economy of data disrupts values, such as privacy, and the way we think about our visibility. In [9], we investigate the digital world from an ethical perspective and a computer science viewpoint. We assess the structure and the dynamic of digital visibility and propose a model-driven approach to handle visibility in service compositions.

5.2. Architecture design for intermediation platforms

During our joint work with Worldline we built a JavaScript compiler for generating dataflow program from plain standard JavaScript sources. In an ACM Middleware conference poster session we raised the question of extracting a dataflow design from JavaScript callback hell. The compiler https://github.com/etnbrd/due-compiler is used to help JavaScript standard developers generate their equivalent dataflow scheme without the need of external libraries such as Promises, Async or Q. With this tool, developer may migrate their javascript legacy code towards a new flow based design. Our due npm module https://github.com/etnbrd/due is a first step towards a dynamic flow based architecture studied in Etienne's project.

The C3PO project provides a browser based application for interacting with other nearby participants in chat mode. The client architecture runs exclusively in the browser over a DTN layer and listens to posts send through a dedicated spontaneous and ephemeral social network (SESN) [5]. The client is organized around a display canvas hosting plugins. Each plugin registers for some tags it wishes to handle. The local DTN manager receives posts and propagates them to the plugins.

We have used intermediation technologies for voting systems. A brief presentation of our motivations has been made in [4]. A patent on the BitBallot protocol is on its way.

DIONYSOS Project-Team

5. New Results

5.1. Highlights of the Year

Pierre L'Ecuyer received the Award of Merit from the Canadian Operational Research Society, 2014.

We had one best paper award in 2014 on a novel architecture for resilient networks (see 5.8).
BEST PAPER AWARD :
[50] **IEEE International Conference on Innovations for Community Services**. D. LEQUÉRÉ, C. BETOULE, G. THOUENON, Y. HADJADJ-AOUL, A. KSENTINI, R. CLAVIER.

5.2. Quality of Experience

Participants: Yassine Hadjadj-Aoul, Adlen Ksentini, Gerardo Rubino, Bruno Sericola, Pantelis Frangoudis, César Viho, Quang Pham Tran Anh.

PSQA. We continue the development of the PSQA technology (Pseudo-Subjective Quality Assessment) in the area of Quality of Experience (QoE). PSQA is today a mature technology allowing to build measuring modules capable of quantifying the quality of a video or an audio sequence, as perceived by the user, when received through an IP network. It provides an accurate and efficiently computed evaluation of quality. Accuracy means that PSQA gives values close to those that can be obtained from a panel of human observers, under a controlled subjective testing experiment, following an appropriate standard (which depends on the type of sequence or application). Efficiency means that our measuring tool can work in real time. Observe that perceived quality is, in general, the main component of QoE when the application or service involves video and audio, or voice. PSQA works by analyzing the networking environment of the communication and some the technical characteristics of the latter. It works without any need to the original sequence (as such, it belongs to the family of *no-reference* techniques). It must be pointed out that a PSQA measuring or monitoring module is network-dependent and application-dependent. Basically, for each specific networking technology, and for any application or service, the module must be built from scratch. But once built, it works automatically and efficiently, allowing if necessary its use in real time, typically for controlling purposes.

QoE and SLA. On the applications side, we focused this year on using QoE estimates to drive service/application-level decisions. As a first use case, we proposed a multi-objective optimization framework for the problem of optimally selecting among a set of available hosting and network connectivity Service-Level Agreements (SLAs) for the migration of enterprise communication services (such as teleconferencing) to the Cloud [59]. Our framework captures the tradeoff between user experience and deployment cost, and offers a service provider the opportunity to weight these two conflicting criteria based on its preferences. Our approach is generic and can be applied to various application settings by appropriately selecting application-specific user experience models. For example, for enterprise voice teleconferencing we used the E-model for estimating user experience under a specific selection of hosting and network SLAs and a specific amount of resources (virtual machines) to deploy.

QoE and collaborative projects. We then considered QoE-aware content delivery, targeting in particular an environment where web and multimedia content is disseminated by over-the-top (OTT) providers, but assuming a level of cooperation between the content provider and the ISP (a trend which has started to become commonplace) [46]. We built on the outcome of our prior work [0], where we designed and implemented a network load estimation methodology and tool which operates by observing the delay behavior of the Precision Time Protocol (PTP) for network clock synchronization. After quantitatively establishing the link between network load and user experience, we proposed an architecture for OTT content delivery where user

[0] P.A. Frangoudis, A. Ksentini, Y. Hadjadj-Aoul, and G. Boime, "PTPv2-based network load estimation," Proc. IEEE ISPCS 2013. (This work was carried out in the context of the FUI project IPChronos, see 6.10 .)

requests are redirected to the data centers expected to offer optimal QoE, taking into account, among others, information about network load in the media path offered by our load estimation service (LES) in real time. In the same context, we developed a demonstrator where the LES is integrated as an additional network probe with the QoE monitoring architecture developed in the Celtic QuEEN project (see 7.2.1.1). Using a simple video QoE model which takes into account network load and video information (quality/resolution, bitrate), we implemented [0] an adaptation scheme for DASH video delivery which switches among video qualities based on QoE estimates received by the QuEEN software agent.

QoE and PTPv2. In [46], we make the case for an alternative use of the PTPv2 protocol: Adopting a learning approach, we observe its delay behavior during the protocol message exchange, derive models of its dependence on network load and build a real-time load estimation service. Then, as an application scenario of this service, we turn our attention to the provision of Over-the-Top (OTT) services. In such an environment, and assuming a level of cooperation between the ISP and the OTT provider, we demonstrate how our service can be used for estimating the QoE for web applications. To this end, we establish quantitatively the link between network load and user experience using a state-of-the-art web QoE monitoring framework, and show how our PTPv2-based load estimation scheme can integrated in an OTT service architecture and be utilized for load-aware, QoE-optimized content delivery decisions.

QoE and reneging. We consider in [45] an important Quality of Experience (QoE) indicator in mobile networks that is reneging of users due to impatience. We specifically consider a cell under heavy load conditions and compute the reneging probability by using a fluid limit analysis. By solving the fixed point equation, we obtain a new QoE perturbation metric quantifying the impact of reneging on the performance of the system. This metric is then used to devise a new pricing scheme accounting for reneging. We specifically propose several flavors of this pricing around the idea of having a flat rate for accessing the network and an elastic price related to the level of QoE perturbation induced by the communications.

QoE-aware OLSR for Video Streaming over Wireless Multihop Networks. Multi-hop environments can impact significantly ad-hoc network performance. In [57], we propose a routing algorithm based on optimized link state routing (OLSR), aimed at guaranteeing the quality of experience (QoE) of users in these types of networks. PSQA (see above in this same section) is used to estimate a mean opinion score (MOS), and then this MOS value is exploited by the source for selecting the appropriate path in the network. Moreover, an event- triggered based on the MOS value is used to provide more relevant information in selecting the best path by the source. The performance of this proposed mechanism was validated through intensive simulation under different scenarios. The results in [57] show that the proposed scheme outperforms other OLSR-based routing protocols particularly in a heavy load and high mobility scenario.

QoE-Aware Routing for Video Streaming over VANETs. In-vehicle multimedia applications are gaining interest since recent years. However, the high loss rate caused by high mobility in vehicular networks (VANETs) imposes several challenges in multimedia transmission. Moreover, in the context of multimedia, the quality of service (QoS)-based approaches assess the quality of streaming services through network-oriented metrics while the concept of quality of experience (QoE) is built upon the perception of users. In [58], a QoE-based routing protocol for video streaming over VANETs is proposed. By taking the mean opinion score (MOS) into account for path selection, good performance levels can be achieved, as shown by our simulation results.

5.3. Analytic models

Participants: Bruno Sericola, Gerardo Rubino, Raymond Marie.

New book about Dependability Theory. Dependability metrics are omnipresent in every engineering field, from simple ones through to more complex measures combining performance and dependability aspects of systems. The new book [69] written in the team, entitled “Markov Chains and Dependability Theory” and published in 2014 by Cambridge University Press (see also http://www.amazon.fr/Markov-Chains-Dependability-Theory-Gerardo/dp/1107007577/), presents the mathematical basis of the analysis of

[0] Our video adaptation scheme is implemented in the VLC open-source media player.

these metrics. The modelling context corresponds to the most used framework, Markov models. The book describes both basic results and specialised techniques. The authors first present discrete and continuous time Markov chains before focusing on dependability measures, which necessitate the study of Markov chains on a subset of states representing different user satisfaction levels for the modelled system. Topics covered include Markovian state lumping, analysis of sojourns on subset of states of Markov chains, analysis of most dependability metrics, fundamentals of performability analysis, and bounding and simulation techniques designed to evaluate dependability measures. As stated in its abstract, the book is of interest to graduate students and researchers in all areas of engineering where the concepts of lifetime, repair duration, availability, reliability and risk are important.

Fluid models. In [77] we study congestion periods in a finite fluid buffer when the input rate depends upon a recurrent Markov process; congestion occurs when the buffer content is equal to the buffer capacity. We consider the duration of congestion periods as well as the associated volume of lost information. We derive their distributions in a typical stationary busy period of the buffer. Our goal is to compute the exact expression of the loss probability in the system, which is usually approximated by the probability that the occupancy of the infinite buffer is greater than the buffer capacity under consideration. Moreover, by using general results of the theory of Markovian arrival processes, we show that the duration of congestion and the volume of lost information have phase-type distributions.

Industrial Logistic Aspects. Motivated by the consideration of clauses of penalty, we worked again on the determination of the probability distributions of the delays of unavailability of systems on the operational sites. By considering in particular a given type of spare, we show the important role played by the possible waiting time of the change during the occurrence of a breakdown. In particular we verify that the cumulative probability distribution of the delay of unavailability possesses a relatively low tail diminution as well as a high square of cœfficient of variation. Upper and lower bounds are highlighted in the simplest case. These results allow to calculate the risk inferred by the use of clauses of penalty; for example, by proposing an expression of the expectation of the cost of penalty imposed by unit of time if any unavailability exceeding a certain threshold is penalized [62]. If the possible waiting time of the change is the obsession of the specialists of the maintenance, the consideration of stock shortages in supply chains is often underestimated when these events are rare events. A related work consisted in showing that a low probability of break can be associated with a high coefficient of variation can have a very significant consequence [54].

We also studied the extension of our analytical method of calculation of the operational availability of a fleet of consequent systems deployed on a site and maintained by exchanges on the site of subsets (the LRU for *line repaired unit*) in the specific case where a policy of cannibalization is implemented. We propose an approximated method which is particularly adapted to the case of systems with strong operational availability because in this case the error inferred by the approximation remains low. The developed method consists in determining the expectation of the number of blocked systems due to the lack of change, in the presence of a policy of cannibalization. This expectation is directly associated with a loss of operational availability. At present, in the presence of a policy of cannibalization, the proposed solution concerns only the systems constituted by a series of LRU but the policy of cannibalization can be applied to all or part of the types of LRU [63].

5.4. Performance Evaluation

Participants: Pierre L'Ecuyer, Bruno Sericola, Romaric Ludinard.

Network Monitoring and Fault Detection. Monitoring a system consists in collecting and analyzing relevant information provided by the monitored devices, so as to be continuously aware of the system state (situational awareness). However, the ever growing complexity and scale of systems makes both real time monitoring and fault detection a quite tedious task. Thus the usually adopted option is to focus solely on a subset of information states, so as to provide coarse-grained indicators. As a consequence, detecting isolated failures or anomalies is a quite challenging issue. We propose in [39], [61] to address this issue by pushing the monitoring task at the edge of the network. We present a peer-to-peer based architecture, which enables nodes to adaptively and efficiently self-organize according to their "health" indicators. By exploiting both temporal and spatial correlations that exist between a device and its vicinity, our approach guarantees that only isolated anomalies (an anomaly is isolated if it impacts solely a monitored device) are reported on the fly to the network operator. We show that the end-to-end detection process, *i.e.*, from the local detection to the management operator reporting, requires a logarithmic number of messages in the size of the network.

Robustness Analysis of Large Scale Distributed Systems. In the continuation of [81] which proposed an in-depth study of the dynamicity and robustness properties of large-scale distributed systems, we analyze in [13], the behavior of a stochastic system composed of several identically distributed, but non independent, discrete-time absorbing Markov chains competing at each instant for a transition. The competition consists in determining at each instant, using a given probability distribution, the only Markov chain allowed to make a transition. We analyze the first time at which one of the Markov chains reaches its absorbing state. When the number of Markov chains goes to infinity, we analyze the asymptotic behavior of the system for an arbitrary probability mass function governing the competition. We give conditions for the existence of the asymptotic distribution and we show how these results apply to cluster-based distributed systems when the competition between the Markov chains is handled by using a geometric distribution.

Detection of distributed deny of service attacks A Deny of Service (DoS) attack tries to progressively take down an Internet resource by flooding it with more requests than it is capable to handle. A Distributed Deny of Service (DDoS) attack is a DoS attack triggered by thousands of machines that have been infected by a malicious software, with as immediate consequence the total shut down of targeted web resources (*e.g.*, e-commerce websites). A solution to detect and to mitigate DDoS attacks it to monitor network traffic at routers and to look for highly frequent signatures that might suggest ongoing attacks. A recent strategy followed by the attackers is to hide their massive flow of requests over a multitude of routes, so that locally, these flows do not appear as frequent, while globally they represent a significant portion of the network traffic. The term "iceberg" has been recently introduced to describe such an attack as only a very small part of the iceberg can be observed from each single router. The approach adopted to defend against such new attacks is to rely on multiple routers that locally monitor their network traffic, and upon detection of potential icebergs, to inform a monitoring server that aggregates all the monitored information to accurately detect icebergs. Now, to prevent the server from being overloaded by all the monitored information, routers continuously keep track of the c (among n) most recent high flows (called items) prior to sending them to the server, and throw away all the items that appear with a small probability. Parameter c is dimensioned so that the frequency at which all the routers send their c last frequent items is low enough to enable the server to aggregate all of them and to trigger a DDoS alarm when needed. This amounts to compute the time needed to collect c distinct items among n frequent ones. A thorough analysis of the time needed to collect c distinct items appears in [71].

Randomized Message-Passing Test-and-Set. In [74], we present a solution to the well-known Test&Set operation in an asynchronous system prone to process crashes. Test&Set is a synchronization operation that, when invoked by a set of processes, returns yes to a unique process and returns no to all the others. Recently, many advances in implementing Test&Set objects have been achieved, but all of them target the shared memory model. In this paper we propose an implementation of a Test&Set object in the message passing model. This implementation can be invoked by any number $p \leq n$ of processes where n is the total number of processes in the system. It has an expected individual step complexity in $O(\log p)$ against an oblivious adversary, and an expected individual message complexity in $O(n)$. The proposed Test&Set object is built atop a new basic building block, called selector, that allows to select a winning group among two groups of processes. We propose a message-passing implementation of the selector whose step complexity is constant. We are not aware of any other implementation of the Test&Set operation in the message passing model.

Call centers. We develop research activities around the analysis and design of call centers, from a performance perspective. In [56], we focus on the scheduling problem (which task must be done by which worker at each period of time). We show that a Constraint Programming model can be used to solve large instances of this type of optimization work. In [21], we study call routing policies for call centers with multiple call types and multiple agent groups, focusing on the case of small and medium size centers, whose behavior may differ from those obtained in heavy-traffic regimes, and for which non-work-conserving policies can perform better. We propose a routing policy based on weights, expressed as linear functions of the call waiting times and agent idle times, or number of idle agents, following a simulation-based optimization approach.

5.5. Network Economics

Participants: Bruno Tuffin, Pierre L'Ecuyer.

The general field of network economics, analyzing the relationships between all acts of the digital economy, has been an important subject for years in the team. The whole problem of network economics, from theory to practice, describing all issues and challenges, is described in our book [67].

Among the topics we have particularly focused on, the network neutrality debate was a major concern in 2014. In the position paper [79], Bruno Tuffin and his co-author Patrick Maillé discuss for a large audience the issues and challenges of network neutrality in response to the European parliament text voted in April 2014. A related (and often forgotten) issue, the recently raised search neutrality debate questions the ranking methods implemented by search engines: when a search is performed, do they (or should they) display the web pages ordered according to the quality-of-experience (relevance) of the content? In [22], we analyze that question in a setting when content is offered for free, content providers making revenue through advertising. For content providers, determining the amount of advertising to add to their content is a crucial strategic decision. Modeling the trade-off between the revenue per visit and the attractiveness, we investigate the interactions among competing content providers as a non-cooperative game, and consider the equilibrium situations to compare the different ranking policies. Our results indicate that when the search engine is not involved with any high-quality content provider, then it is in its best interest to implement a neutral ranking, which also maximizes user perceived quality-of-experience and favors innovation. On the other hand, if the search engine controls some high-quality content, then favoring it in its ranking and adding more advertisement yields a larger revenue. This is not necessarily at the expense of user perceived quality, but drastically reduces the advertising revenues of the other content providers, hence reducing their chances to innovate.

But while ISPs and search engines are almost the only Internet actors being pointed out as potentially non neutral, we investigate the economic impact and strategies of Content Delivery Networks (CDNs), Internet actors that reduce the capacity needs in the backbone network and improve the quality perceived by users. The growing importance of Content Delivery Network (CDN) in the value chain of content delivery raises concerns about the neutrality of these players. We consider in [52] the so-called push and pull models where the traffic is paid by the sender or the receiver, respectively, as well as the situation where the CDN is (vertically) integrated to, i.e., owned by, an Internet Service Provider (ISP). We then discuss the implication of CDNs into the network neutrality debate, another issue forgotten by researchers and regulators. We also propose in [53] a model to analyze the impact of revenue-oriented CDN management policies on the fairness of the competition among two content providers that use CDN services to deliver contents. We show that there exists a unique optimal revenue maximizing policy for a CDN actor –the dimensioning and allocation of its storage capacity– that depends on prices for service/transport/storage, and on the distribution of content popularity. Using data from the analysis of traces from two major content providers (YouTube Live and justin.tv), we remark that a CDN remains a relatively neutral actor even when one of the content providers it serves tries to monopolize the CDN storage space by implementing an aggressive policy to harm its competitors.

Finally, when a customer searches for a keyword at a classified ads website, at an online retailer, or at a search engine (SE), the platform has exponentially many choices in how to sort the output to the query. The two extremes are (a) to consider a ranking based on relevance only, which attracts more customers in the long run because of perceived quality, and (b) to consider a ranking based on the expected revenue to be generated by immediate conversions, which maximizes short-term revenue. Typically, these two objectives are not perfectly positively correlated and hence the main question is what middle ground between them should be chosen. We introduce in [78] stochastic models and propose effective solution methods that can be used to optimize the ranking considering long-term revenues. A key feature of our model is that customers are quality-sensitive and are attracted to the platform or driven away depending on the average relevance of the output. The proposed methods are of crucial importance in e-business and encompass: (i) classified ad websites which can favor paid ads by ranking them higher, (ii) online retailers which can rank products they sell according to buyers' interests and/or the margins these products have, (iii) SEs which can position the content that they serve higher in the output page than third-party content to keep users in their platforms for longer and earn more. This goes in detriment of just offering rankings based on relevance only and is directly linked to the current search neutrality debate.

5.6. Monte Carlo

Participants: Bruno Tuffin, Gerardo Rubino, Pierre L'Ecuyer.

We maintain a research activity in different areas related to dependability, performability and vulnerability analysis of communication systems, using both the Monte Carlo and the Quasi-Monte Carlo approaches to evaluate the relevant metrics. Monte Carlo (and Quasi-Monte Carlo) methods often represent the only tool able to solve complex problems of these types. A review of Monte Carlo, Quasi-Monte Carlo and pseudo-random generation can be found in [66]. In [27], we examine some properties of the points produced by certain classes of long-period linear multiple recursive random number generators. These generators have their parameters selected in special ways to make the implementation faster. We show that as a result, the points produced by these generators have a poor lattice structure, and a poor initialization of the state can have long-lasting impact, because of the limited diffusion capacity of the recurrence.

However, when the events of interest are rare, simulation requires a special attention, to accelerate the occurrence of the event and get unbiased estimators of the event of interest with a sufficiently small relative variance. This is the main problem in the area. Dionysos' work focuses then on dealing with the rare event situation. In [20], we present several state-of-the-art Monte Carlo methods for simulating and estimating rare events. Among variance reduction methods, the most prominent ones for this purpose are Importance Sampling (IS) and Multilevel Splitting, also known as Subset Simulation. Some recent results on both aspects are described, motivated by theoretical issues as well as by applied problems.

A non-negligible part of our activity on the application of rare event simulation was about the evaluation of static network reliability models, with links subject to failures. Exact evaluation of static network reliability parameters belongs to the NP-hard family and Monte Carlo simulation is therefore a relevant tool to provide their estimations. In [34], we propose an adaptive parameterized method to approximate the zero-variance change of measure. The method uses two rough approximations of the unreliability function, conditional on the states of any subset of links being fixed. One of these approximations, based on mincuts, under-estimates the true unknown unreliability, whereas the other one, based on minpaths, over-estimates it. Our proposed change of measure takes a convex linear combination of the two, estimates the optimal (graph-dependent) coefficient in this combination from pilot runs, and uses the resulting conditional unreliability approximation at each step of a dynamic importance sampling algorithm. This new scheme is more general and more flexible than a previously-proposed zero-variance approximation one, based on mincuts only, and which was shown to be robust asymptotically when unreliabilities of individual links decrease toward zero. Our numerical examples show that the new scheme is often more efficient when low unreliability comes from a large number of possible paths connecting the considered nodes rather than from small failure probabilities of the links. Another paper, reference [18], focuses on another technique, known as Recursive Variance Reduction (RVR) estimator which approaches the unreliability by recursively reducing the graph from the random choice of the first working link on selected cuts. This previously known method is shown to not verify the bounded relative error (BRE) property as reliability of individual links goes to one, i.e., the estimator is not robust in general to high reliability of links. We then propose to use the decomposition ideas of the RVR estimator in conjunction with the IS technique. Two new estimators are presented: the first one, called Balanced Recursive Decomposition estimator, chooses the first working link on cuts uniformly, while the second, called Zero-Variance Approximation Recursive Decomposition estimator, combines RVR and our zero-variance IS approximation. We show that in both cases BRE property is verified and, moreover, that a vanishing relative error (VRE) property can be obtained for the Zero-Variance Approximation RVR under specific sufficient conditions. A numerical illustration of the power of the methods is provided on several benchmark networks. Continuing the analysis of existing method, we have described in [44] a necessary and sufficient condition for a well known technique called Fishman's method to verify BRE and have realized a deep analysis of the technique.

But in the literature and the previously described static network reliability models one typically assumes that the failures of the components of the network are independent.This simplifying assumption makes it possible to estimate the network reliability efficiently via specialized Monte Carlo algorithms. Hence, a natural question to consider is whether this independence assumption can be relaxed, while still attaining an elegant and tractable model that permits an efficient Monte Carlo algorithm for unreliability estimation. In [75], we provide one possible answer by considering a static network reliability model with dependent link failures, based on a Marshall-Olkin copula, which models the dependence via shocks that take down subsets of components at exponential times, and propose a collection of adapted versions of permutation Monte Carlo (PMC, a conditional Monte Carlo method), its refinement called the turnip method, and generalized splitting (GS) methods, to estimate very small unreliabilities accurately under this model. The PMC and turnip estimators have bounded relative error when the network topology is fixed while the link failure probabilities converge to 0. When the network (or the number of shocks) becomes too large, PMC and turnip eventually fail, but GS works nicely for very large networks, with over 5000 shocks in our examples. [65] focuses on the application of our zero-variance approximation IS estimator to this same type of model.

Another family of models of interest in the group are the highly reliable Markovian systems, made of components subject to failures and repairs. We describe in [60] how importance sampling can be applied to efficiently estimate the average interval availability of those models. We provide a methodology for approximating the zero-variance change of measure. The method is illustrated to be very efficient on a small example, compared with standard importance sampling strategies developed in the literature.

Finally, in Quasi-Monte Carlo (QMC), the error when estimating an integral uses a deterministic sequence (instead of a random one) called a low discrepancy sequence and having the property to spread quickly over the integration domain. The estimation error is bounded by the product of a quantity depending on the discrepancy of the sequence and the variation of the integrand. But this bound is proved to be useless in practice. By combining MC and QMC methods, we can benefit from the advantages of both approaches: error estimation from MC and convergence speed from QMC. Randomized quasi-Monte Carlo (RQMC) is another class of methods for reducing the noise of simulation estimators, by sampling more evenly than with standard MC. In [37], we analyze the convergence rate of the *array-RQMC* technique, a randomized QMC method we have previously designed and devoted to the simulation of Markov chains.

In [19], we propose a method for estimating performability metrics built upon non-binary network states, determined by the hop distances between distinguished nodes. In other words, we explore the analysis of a generalization of network reliability, particularly relevant for instance in telecommunications. The estimation is performed by a Monte Carlo simulation method where the sampling space is reduced using edge sets known as d-pathsets and d-cutsets. Numerical experiments over two mesh-like networks are presented. They show significant efficiency improvements relative to the crude Monte Carlo method, in particular as link failures become rare events, which is usually the case in most real communication networks.

5.7. Wireless Networks

Participants: Osama Arouk, Btissam Er-Rahmadi, Adlen Ksentini, Yassine Hadjadj-Aoul, Quang Pham Tran Anh, Hyunhee Park, César Viho.

We continue our activities around wireless and mobile networks, where we focus particularly on 4G/5G networks as well as on a new mobile architecture known as mobile cloud.

LTE improvements. In [35], we investigated, at both the core network (EPC) and Radio Access Network (RAN), the impact of caching the shared content among users. We reviewed the different locations were data could be cached and their impacts on user QoS/QoE. In [33], we proposed several new mechanisms to handle the gateway relocation in the context of highly decentralized mobile network. To evaluate these mechanisms, we proposed an analytical model based on Markov Chains, whereby we captured the randomness of user mobility and its impact on the user QoS in terms of the probability to be connected to the optimal gateway, the drop rate, etc. In [32], we devised an agile admission control mechanism that anticipates QoS/QoE degradation and proactively defines policies for admitting UEs handing-in from the macro network to the small cell network. It also enables IP flow mobility between small cells and macro networks. We provided an analytical model to the admission control mechanism based on Markov Decision Processes (MDP). The ultimate objective of the proposed model is to derive the optimal policy (i.e., reject or accept flows in the macro or the small cell) which maximizes users' QoE under different load scenarios (low and high load user traffic). Another work regarding small cells in LTE was proposed in [76], where we used the small cell principle to extend the mobile network coverage in emerging countries that not include a wired infrastructure. The proposed framework aims to backhaul the small cell with the less costly connection, while ensuring minimal QoS to users. In this vein, we formulated this problem through an Integer Linear Program (ILP), and solve it for small network sizes. For large instances of the network size, we proposed two new heuristics. In [30], we investigated network decentralization in conjunction with the Selective IP Traffic Offload (SIPTO) approaches to handle the mobile increased data traffic. We first devised different approaches based on a per destination domain name basis, which offer operators a fine-grained control to determine whether a new IP connection should be offloaded or accommodated via the core network. Two of our solutions are based on Network Address Translation (NAT) named simple-NATing and twice-NATing, while a third one employs simple tunneling and a forth proposal adopts multiple Access Point Names (APNs). We also proposed methods enabling User Equipment (UEs), both in idle and active modes and while being on the move, to always have efficient Packet Data Network (PDN) connections. A qualitative analysis and a simulation study compared the different approaches with respect to cost, complexity, service continuity and network performance, demonstrating the significance of the proposed schemes for multimedia applications.

M2M. We addressed another type of traffic that appeared these last years, namely Machine to Machine (M2M) communication or Machine Type Communication (MTC). Such traffic is known by its intensity and its impact on increasing congestion in both parts of 4G networks, the Radio Access Network (RAN) and the core network. The main spirit of the proposed solutions is to proactively anticipate system overload by reducing the amount of MTC signaling messages exchanged in normal network operations. In [49] we introduced a solution that operates at the core network. We proposed that the Mobility Management Entity (MME), or an alike core network node, computes the device trigger rate that alleviates congestion, and communicates this value to the MTC-Interworking Function (MTC-IWF) element that enforces MTC traffic control, via admission control or data aggregation, on the device trigger request rate received from the different MTC servers.

As mentioned earlier, the MTC would impact not only the EPC part, but also the RAN. Group paging is currently considered as one of the most efficient mechanisms proposed to alleviate the problem of the RAN overload. In [42] we introduced a new solution to improve the performance of the current group paging method and overcome its disadvantages. The proposed solution is intended for MTC devices in connected mode state, in which they have an RRC context without being synchronized with the network. In [41] we devised a novel algorithm which estimates the network status (the number of active devices), thus better controlling the RAN access. Unlike most existing methods that consider only one channel, the proposed solution uses the statistics of all the channels in order to estimate the number of arrivals (UE and MTC devices) in each RA (Random Access) slot.

Most of the above-proposed solutions are basically incremental ones. In [31], we devised a complete new architectural vision to support MTC in mobile networks. This vision relies on the marriage of mobile networks and the cloud, specifically based on Network Function Virtualization (NFV). The proposed solution simplifies the network attach procedure for MTC devices by creating only one NFV MTC function that groups all the usual procedures. By doing so, the proposed solution is able to create and scale instances of NFV MTC functions on demand and in an elastic manner to cope with any sudden increase in traffic generated by MTC devices.

Wireless Sensor Networks (WSN). WSNs are complex systems that are mainly limited by the battery life of the nodes in order to have an adequate performance. In most cases, it is possible to have a re-deployment of new nodes in order to prolong the systems lifetime. This leads to a situation where some nodes have a low energy level while other nodes (the majority of nodes a few instants after the re-deployment procedure) have high energy levels. In these environments, it is clear that ancient nodes, those with low energy levels, have to contend for the shared medium against the majority of high energy nodes. As such, the remaining battery life of low energy nodes would be rapidly consumed. In [64], we propose to extend the battery life of low energy nodes by means of assigning prioritized access to the shared channel to those nodes. The goal is to content among a low population of such nodes, while delaying the contention access of high energy nodes which can support higher number of collisions before energy depletion. This is done by studying two different transmission strategies referred to as “hard” and “soft” transmission probabilities. Results show that a soft transmission strategy achieves better results in terms of reduced energy consumption than both the conventional protocol or a hard transmission assignment.

The communication between nodes is the greedy factor to the energy consumption. One important mechanism to reduce the energy consumption is the in-network data aggregation. This mechanism removes repeated and unnecessary data readings and thus cuts on the energy used in communications. In [14] we reviewed the state of art on this topic. Then, we proposed a classification of the available solutions according to the way the aggregation is done. In [15], we addressed the reliable minimum data aggregation scheduling problem in wireless sensor networks under multi-channel frequency use. The proposed solution ensures the collection reliability and reduces the latency in disseminating aggregated data to the base-station over multi-frequency radio links. Another mechanism to improve energy efficiency is to optimize link scheduling when using TDMA-based techniques and data fragmentation when using slotted CSMA/CA access methods. In this line, we proposed a protocol, named DLSP, with the objective of achieving both low energy consumption and low latency in Wireless Sensor Networks. DLSP takes advantage of the spatial reuse of interference-free time slots by means of conflicts graphs. Unlike the previous studies that often consider saturated nodes, we propose to relax the saturation assumption in order to maintain good performance when some of the nodes have no data to send citemouloud:hal-01101396. In [55], we noticed that the standardized slotted CSMA/CA may lead to a wastage of the bandwidth utilization and an additional transmission delay. This drawback is mainly caused by Deferred Transmission in the CSMA/CA algorithm at the end of the superframe, when there is not sufficient time to complete the frame transmission. Thus, we proposed to fragment a data frame into a short frame and attempt its transmission in the current frame and transmit the remaining frame in the next superframe. The data fragmentation mechanism was modeled using a Markov chain. A non-saturated traffic and acknowledgement transmission are considered in our analysis.

High data rate WiFi networks. The IEEE 802.11ac Task Group (TGac) is actively working on an amendment that allows WLAN to reach a maximum aggregate network throughput up to 7 Gbps on bands below 6 GHz. In particular, the standard envisions a maximum Medium Access Control (MAC) throughput of at least 500 Mbps for a single user, and at least 1 Gbps in case of multiple users. In [36] we proposed an analysis of the IEEE 802.11ac TXOP Sharing mechanism, which was recently introduced by the 802.11ac group, by providing a Markov chain-based model. Based on the proposed Markov chain, we provided an analytical model of the achievable throughput for each AC. Accordingly, we can analyze the impact of the TXOP Sharing on the throughput of each AC, hence highlighting the improvement achieved in terms of bandwidth utilization and channel access fairness among the different ACs.

Mobile cloud. One of the 5G-architecture visions considers the usage of clouds to build mobile networks and help in decentralizing mobile networks on demand, elastically, and in the most cost-efficient way. This concept of carrier cloud becomes of vital importance knowing that several cloud providers are distributing their cloud/network, globally deploying more regional data centers, to meet their ever-increasing business demands. As an important enabler of the carrier cloud concept, network function virtualization (NFV) is gaining great momentum among industries. NFV aims for decoupling the software part from the hardware part of a carrier network node, traditionally referring to a dedicated hardware, single service and single-tenant box, that is using virtual hardware abstraction. Network functions become thus a mere code, runnable on a particular, preferably any, operating system and on top of a dedicated hardware platform. The ultimate objective is to run network functions as software in standard virtual machines (VMs) on top of a virtualization platform in a general-purpose multi-service multi-tenant node (e.g., Carrier Grade Blade Server) put into the cloud. In [26], we presented a LISP-based implementation of the Follow Me Cloud (FMC) concept, whereby mobile services hosted in federated clouds follow mobile users as they move and according to their needs. This implementation clearly demonstrates the feasibility of the FMC concept. On the other hand, service migration in FMC may be an expensive operation given the incurred cost in terms of signaling messages and data transferred between DCs. Indeed, decision on service migration defines therefore a tradeoff between cost and user perceived quality. In [48] we addressed this tradeoff by modeling the service migration procedure using a Markov Decision Process (MDP). The aim was to formulate a decision policy that determines whether to migrate a service or not when the concerned User Equipment (UE) is at a certain distance from the source DC.

In order to meet the general needs of mobile operators, efficient mobile cloud must give high importance to the placement/instantiation of mobile network functions (such as data anchor gateways) in the federated cloud. In [43] we argued the need of using service/application type and requirements as metrics for efficiently: (i) create virtual instance of the Packet Data Network Gateway (PDN-GW); (ii) select the virtual PDN-GW for UEs with specific application type. After modeling this procedure though a nonlinear Optimization Problem (OP) and proving it as a NP-hard problem, we proposed three solutions to solve this issue.

Wireless Local Area Networks. User-centric networking has emerged as a disruptive new communication paradigm. We particularly focused on its expressions in *wireless* networking and the challenges it brings about [23]. In this context, by means of testbed experiments and simple analytic models, we quantified the upper bounds on VoIP capacity of a purely user-centric secure VoIP communications scheme that we designed, identifying the major quality degradation factors. Our results have shown that typical user Wi-Fi equipment can sustain a satisfactory number of concurrent secure VoIP sessions with acceptable QoE and, at the same time, protection from malicious user activity can be offered to access providers, while a level of roaming privacy can be guaranteed [24]. We then studied the role of users in wireless network management tasks. In particular, we proposed a scheme where monitoring the topology of Wi-Fi deployments is crowdsourced to roaming users, who submit reports on wireless coverage in their vicinity [25]. Topology information can then be used as input to reconfiguration mechanisms, such as channel assignment schemes. Users cannot be assumed trustworthy, though. They can engage in fraudulent reporting, which, unless specific countermeasures are in place, can severely impact one's view of the network topology. To this end, we designed and implemented an architecture for accurate Wi-Fi topology discovery, devising a reputation-based mechanism to tackle realistic and simple to implement attacks. We have shown analytically and via simulation that, even in the presence of large numbers of attackers, our user-centric scheme significantly outperforms pure infrastructure-based approaches, where monitoring is carried out only by trusted Access Points.

In another line of research, we focused on efficiently integrating wireless users in an Information-Centric Network (ICN) architecture. In ICN, multicast content delivery is the norm. At the same time, wireless multicast is problematic. To address this issue, we took advantage of the content awareness inherent in ICN and proposed a relay-based approach for local wireless multicasting: ICN information *scoping* mechanisms assist in expressing content semantics and, in turn, encoding the heterogeneous performance requirements of different content/application types. Under this premise, we proposed a multiobjective optimization approach for relay selection and multicast transmission rate assignment which allows to optimize for reliability, delivery time, or energy cost on a per content basis [47].

Energy saving. Another part of our activities in wireless network are related to energy saving. Indeed, one of the biggest problem today in the wireless world is that wireless devices are battery-driven, which reduce their operating lifetime. The experimental measurements we have achieved in [16] and [17] revealed that operating system overhead causes a drop in performance and energy consumption properties as compared to the GPP in case of certain low video qualities. We propose, thus, a new approach for energy-aware processor switching (GPP or DSP) which takes into consideration video quality. We show the pertinence of our solution in the context of adaptive video decoding and implement it on an embedded Linux operating system.

Adaptive Beam Scheduling for Scalable Video Multicast in Wireless Networks. Design of efficient multicast for a scalable video coding (SVC) streaming combined with directional beamforming is a challenging issue. In [29], we propose a QoE-aware directional beam scheduling (QBS) scheme which optimizes overall quality of experience (QoE) for multirate multicast of SVC, with beamforming in wireless networks. We optimally schedule different SVC layers to different beams and rate modulations. We provide a mixed integer linear programing (MILP) formulation of the problem, and then propose a heuristic algorithm. Extensive simulation results demonstrate that QBS can increase the overall QoE and can satisfy a minimum expected QoE for all users.

5.8. Future networks and architectures

Participants: Damien Le Quéré, Adlen Ksentini, Yassine Hadjadj-Aoul, Jean-Michel Sanner.

LOCARN. LOCARN (i.e. Low Opex & Capex Architecture for Resilient Networks) is a flat, dynamic and very simple packet architecture that focuses on plug-and-play guidance to provide flexibility and resiliency on the transport of client data traffics. To that end, the counterpart of the solution is a significant overhead due to the generation of control plane packets. In [50], we proved that in typical meshed operators transport networks applications, (i.e. infrastructures having high data-rates and high resiliency requirements), the LOCARN overhead is acceptable up to thousands of communications. In [51], we introduced two proposals that permit to increase the amount of simultaneous communications while maintaining the good properties of the initial design.

SDN. We started an activity on Software Defined Networking (SDN), a recent idea proposed to handle network management problems. SDN are becoming an important issue with the ever-increasing network complexity. They are proposed as an alternative to the current architecture of the Internet, which cannot meet the supported services requirements such as Quality of Service/Experience (Qos/QoE), security and energy consumption. We particularly address the scalability issue by proposing an automated hierarchical controller-based architecture handling the whole control chain.

DISCO Project-Team

6. New Results

6.1. Zero-parameter mono and multi objective methods for the tuning of controllers

The synthesis of controllers for any kind of system is the main point in Automatic Control. The traditional approach is to use a simplified model of the system to control and/or use some reformulations of the specifications to tune an often efficient but suboptimal controller. In a more and more competitive industrial context, the design of high performances controllers has emerged as a crucial point to enhance the global productivity. However, the design of optimal controllers supposes the solution of non-convex and non-differentiable optimization problems, for which deterministic and (often) local search algorithm fail in the solution. In this work, Particle Swarm Optimization is used to solve the problem, and tested to define some controllers for a magnetic levitation. The use of standard settings and penalization terms leads to a zero-parameter and reformulation free method. Results are much than satisfactory and show that Evolutionary Computation could be of great interest in the Automatic Control field.

6.2. Fixed-structure H_∞ synthesis for multiple plants

This work proposes an efficient evolutionary approach to the fixed-order and structured H_∞ control design problem extended to the multiple plants case. By testing it on the classical example of a flexible plant, this evolutionary approach proves to be very efficient compared with other recent tools, especially in the case of a high number of plants; it can then be considered as an interesting alternative for such problems.

6.3. Fixed-Structure Mu-Synthesis

This work proposes to shed a new light on the Mu-synthesis problem using the differential evolution algorithm. This algorithm allows optimizing simultaneously the structured controller and the dynamic (or static) D-scalings, which leads to robust performance controllers. This method has been applied successfully to a classical flexible plant control problem. After a comparison between the evolutionary approach and the non-smooth optimization one has envisaged proving the high potential of the proposed method.

6.4. Algebraic Analysis Approach to Linear Functional Systems

6.4.1. Serre's reduction problem

The purpose of this work is to study the connections existing between Serre's reduction of linear functional systems - which aims at finding an equivalent system defined by fewer equations and fewer unknowns - and the decomposition problem - which aims at finding an equivalent system having a diagonal block structure - in which one of the diagonal blocks is assumed to be the identity matrix. In order to do that, in [62], we further develop results on Serre's reduction problem and on the decomposition problem. Finally, we show how these techniques can be used to analyze the decomposability problem of standard linear systems of partial differential equations studied in hydrodynamics such as Stokes equations, Oseen equations and the movement of an incompressible fluid rotating with a small velocity around a vertical axis.

6.4.2. A spectral sequence central in the behaviour approach

Within the algebraic analysis approach to multidimensional systems, the behavioural approach developed by J. C. Willems can be understood as a dual approach to the module-theoretic approach. This duality is exact when the signal space is an injective cogenerator module over the ring of differential operators. In particular, the obstruction to the existence of a parametrization of a multidimensional system is characterized by the existence of autonomous elements of the multidimensional system. In [52], we consider the case of a general signal space and investigate the connection between the algebraic properties of the differential module defining the multidimensional system and the obstruction to the existence of parametrizations of the multidimensional system. To do so, we investigate a certain Grothendieck spectral sequence connecting the obstructions to the existence of parametrizations to the obstructions to the differential module - defining the multidimensional system - to be torsion-free, reflexive ...projective.

6.4.3. Restrictions of n-D systems and inverse images of D-modules

The problem of characterizing the restriction of the solutions of a n-D system to a subvector space of $\mathbb{R}^n$ has recently been investigated in the literature of multidimensional systems theory. For instance, this problem plays an important role in the stability analysis and in stabilization problems of multidimensional systems. In this work, we characterize the restriction of a n-D behaviour to an algebraic or analytic submanifold of $\mathbb{R}^n$. In [51], within the algebraic analysis approach to multidimensional systems, we show that the restriction of a n-D behaviour to an algebraic or analytic submanifold can be characterized in terms of the inverse image of the differential module defining the behaviour. Explicit characterization of inverse images of differential modules is investigated. Finally, we explain Kashiwara's extension of the Cauchy-Kowalevsky theorem for general n-D behaviours and non-characteristic algebraic or analytic submanifolds.

6.4.4. Artstein's transformation of linear time-delay systems

Artstein's classical results show that a linear first-order differential time-delay system with delays in the input is equivalent to a linear first-order differential system without delays thanks to an invertible transform which includes integral and delay operators. Within a constructive algebraic approach, we show how Artstein's reduction can be found again and generalized as a particular isomorphism problem between the finitely presented modules defined by the two above linear systems over the ring of integro-differential time-delay operators. Moreover, we show that Artstein's reduction can be obtained in an automatic way by means of symbolic computation, and thus can be implemented in computer algebra systems.

6.5. New Techniques for Robust Control of Linear Infinite-Dimensional Systems

6.5.1. Robust stabilization of a flexible rod moving in rotation and translation

We develop a hierarchy of models for a flexible rod moving in rotation and translation from a nonlinear partial differential model (generalization of the Euler-Bernoulli equation) to a linear partial differential equation and finite-dimensional models via approximations. We study the stability of those models as well as their robust stabilizations. This work is an extension of the results obtained in [61]. This work will be pursued within the framework of a CIFRE PhD thesis developed in collaboration with SAGEM (2015).

6.5.2. Noncommutative geometric approach to infinite-dimensional systems theory:

This new field of research aims at showing that noncommutative geometric structures such as connections and curvatures exist on internally stabilizable infinite-dimensional linear systems and on their stabilizing controllers. To see this new geometry, using the noncommutative geometry developed by Connes, we have to replace the standard differential calculus by the quantized differential calculus and classical vector bundles by projective modules. In [50], we give an explicit description of the connections on an internally stabilizable system and on its stabilizing controllers in terms of the projectors of the closed-loop system classically used in robust control. Their curvatures are explicitly computed. These connections aim at studying the variations of the signals in the closed-loop system in response to a disturbance or a change of the reference. The study of these connections are useful to understand how techniques of (noncommutative) differential geometry can be used in the study of H^∞ control theory.

6.5.3. A fractional ideal approach to the robust regulation problem

We show how fractional ideal techniques developed in [8] can be used to obtain a general formulation of the internal model principle for stabilizable infinite-dimensional SISO plants which do not necessarily admit coprime factorization. This result is then used to obtain necessary and sufficient conditions for the robust regulation problem. In particular, we find again all the standard results obtained in the literature.

6.5.4. Robust control as an application to the homological perturbation lemma:

Within the lattice approach to transfer matrices developed in [8], we have recently shown how standard results on robust control can be obtained in a unified way and generalized when interpreted as a particular case of the so-called Homological Perturbation Lemma. This lemma plays a significant role in algebraic topology, homological algebra, algebraic and differential geometry, computer algebra, ... Our results show that it is also central to robust control theory for infinite-dimensional linear systems.

6.6. Set invariance for discrete-time delay systems

We studied the existence of positively invariant sets for linear delay-difference equations. In particular, we regarded two strong stability notions: robust (with respect to delay parameter) asymptotic stability for the discrete-time case and delay-independent stability for the continuous-time case. The correlation between these stability concepts is also considered. Furthermore, for the delay-difference equations with two delay parameters, we provided a computationally efficient numerical routine which is necessary to guarantee the existence of contractive sets of Lyapunov–Razumikhin type. This condition also appears to be necessary and sufficient for the delay-independent stability and sufficient for the robust asymptotic stability. The results are published in [25].

We proposed a new perspective on the structural properties of invaraint sets for time delay systems via set factorization. This novel perspective describes, in a unified framework, different existing notions of invariant sets [60]. Additionally, it is shown that the (possible non-minimal) state space representation is a key element in the description of low complexity invariant sets.

6.7. Low complexity constrained control

On one side, we proposed an explicit (piecewise affine feedback) control obtained via interpolation for constrained linear systems [23]. On another side, we studied the Linear Constrained Regulation problem for Continuous-Time Systems in the presence of non-convex constraints [32]. This might prove to be useful for the multi-agent dynamical systems operating under collision avoidance constraints.

6.8. Fault detection based on set theoretic methods and connexions with fault tolerant control

We proposed a set-theoretic fault detection mechanism for multisensor systems with a classification of possible functioning according to the use in the feedback mechanism. The healthy, faulty and under-recovery class are characterized via set descriptions in the residual space and as such can be monitored via on-line mechanisms [26]. Furthermore, the robust detection has been enhanced with an interval observer mechanism for the monitorig during the transients [28].

6.9. Interval Observer

We made several progresses in the domain of the construction of state estimators called interval observers.

1) In [16], we have shown how interval observers can be constructed for nonlinear (and not Lipschitz) systems possessing a special triangular system. These systems are not cooperative and not globally Lipschitz and have a rather general structure which may result from a change of coordinates or an output injection. Besides, under additional assumptions, input to state stability (ISS) properties are derived. We illustrated the constructions by designing a framer and an ISS interval observer for two models of bioreactors.

2) The contributions [17] and [18] present major results for the design of interval observers for discete-time systems. In [18], coordinate transformations which change an arbitrary linear discrete-time system into a positive one and general nonlinear designs of interval observers for nonlinear systems (satisfying a restricitive stability assumption) are proposed. In [17], it is explained how two classical Luenberger observers can be used, (even in the absence of the positivity property of the studied system or the error equations) as interval observer, provided two appropriate outputs, which compose the lower and the upper bound of the interval observer, are selected.

3) In [33], we present a new type of interval observers for nonlinear systems that are affine in the unmeasured part of the state.They are composed of two copies of classical observers and upper and lower bounds which are designed by taking advantage of positivity properties of the error equations when written in appropriate coordinates.

6.10. Reduction model approach: new advances

We solved several distinct problems entailing to the celebrated reduction model approach. Let us recall that this technique makes it possible to stabilize systems with arbitrarily large pointwise or distributed delays in the input.

1) In [46], solutions to the problem of globally exponentially stabilizing linear systems with an arbitrarily long pointwise delay with sampled feedbacks are presented. The main result of a contribution by F. Mazenc and D. Normand-Cyrot is recalled and compared with other results available in the literature.

2) We considered in [41] a stabilization problem for continuous-time linear systems with discrete-time measurements and a sampled input with a pointwise constant delay. In a first step, we constructed a continuous-discrete observer which converges when the maximum time interval between two consecutive measurements is sufficiently small. We also constructed a dynamic output feedback through a technique which is strongly reminiscent of the reduction model approach. It stabilizes the system when the maximal time between two consecutive sampling instants is sufficiently small. No limitation on the size of the delay was imposed.

3) In [43], we studied a general class of nonlinear systems with input delays of arbitrary size. We adapted the reduction model approach to prove local asymptotic stability of the closed loop input delayed systems, using feedbacks that may be nonlinear. We determined estimates of the basins of attraction for the closed loop systems using Lyapunov-Krasovskii functionals.

4) The contribution [21] is devoted to stabilization problems for time-varying linear systems with constant input delays. The reduction model approach we proposed ensures a robustness property (input-to-state stability) with respect to additive uncertainties, under arbitrarily long delays. It applies to rapidly time-varying systems, and gives a lower bound on the admissible rapidness parameters. We also covered slowly time-varying systems, including upper bounds on the allowable slowness parameters. We illustrated our work using a pendulum model.

6.11. Nonlinear systems with delay

1) In [45], we developed a new technique for stability analysis for nonlinear dynamical systems with delays and possible discontinuities. In contrast with Lyapunov based approaches, the trajectory based approach we proposed involves verifying certain inequalities along solutions of auxiliary systems. It applies to a wide range of systems, notably time-varying systems with time-varying delay, ODE coupled with difference equations, and networked control systems with delay. It relies on the input-to-state stability notion, and yields input-to-state stability with respect to uncertainty.

2) In [39], to address various types of delays including the neutral-type arising in dynamical networks, we dealt with coupled delay differential and continuous-time difference equations and proposed stability and robustness criteria. In these criteria, differential equation parts do not necessarily exhibit unbounded dissipation rate. Subsystems described by differential equations are not required to be input-to-state stable either. No assumptions on network topology are made. To handle such a general case, we construct explicit Lyapunov-type functionals. We established stability and robustness of the overall networks.

3) In [22], [42] and [44], stability results for several families of systems with delay are established. The key ingredient of these contributions is the use of comparison systems of a new type, the theory of the positive systems and linear Lyapunov functionals. We provided robustness of the stability with respect to multiplicative uncertainty in the vector fields. We allowed cases where the delay may be unknown, and where the vector fields defining the systems are not necessarily bounded. We illustrate our work using a chain of integrators and other examples.

6.12. Strictification

In [40], the problem of stabilizing rigid-body attitude dynamics in the presence of pointwise time-delay for the input torque is considered. A quaternion-based linear state feedback controller is shown to achieve local stability in addition to the characterization of sufficient condition that depends only on the magnitude of the initial angular rates. More specifically, no restrictions are imposed on the body initial orientation which is a significant contrast with other results from recent literature that adopt three-dimensional representations for the attitude kinematics. Using the quaternion-based linear feedback structure, the closed-loop system is shown to never admit the possibility for finite-time escapes. While the actual magnitude of the time-delay can be unknown, an upper bound on the delay is assumed to be known. The proof relies on the construction of a functional which does not belong to the family of the strict Lyapunov-Krasovskii functionals, but shares important features with the functionals of this family. The stability conditions and results are illustrated through numerical simulations.

6.13. Stability analysis of fractional and classical neutral systems with commensurate delays

Fractional and classical neutral systems with commensurate delays have chains of poles asymptotic to vertical lines (see [66] for classical systems). The delicate case where system have some chains of poles asymptotic to the imaginary axis is interesting as the absence of poles in the open left half-plane does not guarantee the H_∞-stability of the system.

Stability analysis of classical or fractional neutral systems with one single chain of poles asymptotic to the imaginary axis has been investigated in [88], [70], [2], [69], where the asymptotic location of poles of neutral chains was given and necessary and sufficient conditions for H_∞-stability were derived.

We have performed a full analysis of classical and fractional systems with multiple chains of poles approaching a set of points on the imaginary axis. Moreover, a unified method to analyze the stability of fractional and classical systems has been derived.

6.14. Stabilization of fractional neutral systems with commensurate delays

We consider strictly proper fractional neutral systems with one delay and one chain of poles asymptotic to the imaginary axis including the case where this chain may approach the axis from the right side. Thus the system may possess infinitely many poles in the right half-plane. For these systems, a Youla-Kučera parametrization regarding H_∞-stability of all stabilizing controllers has been obtained in [59]. Having in mind the robustness of the closed-loop relative to parameter uncertainties, we wish to find controllers which are able to provide a closed-loop free of chain of poles asymptotic to the imaginary axis. However, we prove that a large class of realizable stabilizing controllers cannot achieve this. [47].

6.15. Stabilization of MISO fractional systems with I/O delays

In order to yield the set of all the stabilizing controllers of a class of MISO fractional systems with delays by mean of Youla-Kučera parametrization regarding H_∞-stability, we are interested in determining coprime factorizations of the transfer function. Explicit expressions of left coprime factorizations and left Bézout factors have been derived in [85]. Explicit expressions of right coprime factorizations and right Bézout factors for some classes of systems have also been derived in [86]. Recently, we obtain right factors for a more general class of systems. Furthermore, we present these right factors in the minimal form, i.e. factors with the minimal number of coefficients to be determined and with the lowest degree. We also obtain left factors in the minimal form.

6.16. Modeling and control of Acute Myeloid Leukemia

Starting from a PDE model of hematopoiesis given in [64], we have derived several models of healthy or cancer cell dynamics in hematopoiesis and performed several stability analyses.

We have proposed in [58] a new mathematical model of the cell dynamics in Acute Myeloid Leukemia (AML) which takes into account the four different phases of the proliferating compartment as well the fast self-renewal phenomenon frequently observed in AML. As was the case in [58] this model is transformed into a distributed delay system and was analyzed here with input-output techniques. Local stability conditions for an equilibrium point of interest are derived in terms of a set of inequalities involving the parameters of the mathematical model.

We have also studied a coupled delay model for healthy and cancer cell dynamics in Acute Myeloid Leukemia consisting of two stages of maturation for cancer cells and three stages of maturation for healthy cells. For a particular healthy equilibrium point, locally stability conditions involving the parameters of the mathematical model have been obtained [30], [31].

We have performed in [29] a stability analysis of both the PDE model of healthy hematopoiesis and a coupled PDE model of healthy and cancer cell dynamics. The stability conditions obtained here in the time domain strengthen the idea that fast self-renewal plays an important role in AML.

A time-domain stability analysis by means of Lyapunov-Krasovskii functionals has been performed on the delay system modeling healthy hematopoiesis for a strictly positive equilibrium point of interest.

6.17. Algebraic geometry techniques for polynomial systems

6.17.1. Testing the structural stability of N-d discrete linear systems

The goal of this work is to propose new computer algebra based methods for testing the structural stability of N-d discrete linear systems. Recall that a discrete linear system given by its transfer function $G(z_1, ..., z_n) = N(z_1, ..., z_n)/D(z_1, ..., z_n)$ is said to be stable if and only if the denominator $D(z_1, ..., z_n)$ is devoid from zero inside the unit complex poly-disc. This fundamental problem in the analysis of N-d systems has been extensively studied these last decades. At the end of the seventies, DeCarlo et al [77] show that testing the previous condition is equivalent to testing the existence of complex zeros on each face of the poly-disc i.e. $D(1, ..., z_i, ..., 1)$ for $i = 1...n$ as well as testing the existence of complex zero on the poly-circle i.e. the zeros of $D(z_1, ..., z_n)$ when $|z_1| = ... = |z_n| = 1$.

Starting from the conditions of DeCarlo et al, we propose a new approach that transform the last condition, that is, the non-existence of complex zeros on the unit poly-circle to a condition on the existence of real solutions inside a region of R^n. More precisely we propose two type of transformations. The first one reduces the problem to looking for real solutions inside the unit box while the second one reduces the problem to looking for real solutions in the whole space R^n. In order to check the existence of real solutions, we use classical computer algebra algorithms for solving systems of polynomial equations. In the case of one or two variables, the appearing systems are generally zero-dimensional. To count or locate the real solutions of such systems, we compute a rational univariate representation [95], that is a one to one mapping between the solutions of the system and the roots of a univariate polynomial, thus the problem is reduced to a univariate problem. When the number of variables is larger than two, the systems that stem from the conditions above are no longer zero-dimensional. In such case, we use critical points method that allow to compute solutions in each real connected component of the zeros of the systems [65].

We implemented the previous approach on maple using the external library *R*aglib [63] which provides routines for testing the existence of real solutions of an algebraic system. Preliminary tests show the relevance of our approach.

This work is supported by the ANR MSDOS grant.

6.17.2. Efficient algorihtms for solving bivariate algebraic systems

This work addresses the problem of solving a bivariate algebraic system (i.e computing certified numerical approximation of the solutions) via the computation or a rational univariate representation. Such a representation is useful since it allows to turn many queries on the system into queries on univariate polynomials. Given two coprime polynomials P and Q in $Z[x,y]$ of degree bounded by d and bitsize bounded by τ we present new algorithms for computing rational univariate representation of the system $\{P,Q\}$ and from this representation, isolating the real solutions of $\{P,Q\}$. The cost analysis of these algorithms show that they have a worst-case bit complexity in $sOB(d^6+d^5\tau)$ which improves by a factor d the state-of-the-art complexity.

DIVERSE Project-Team

6. New Results

6.1. Highlights of the Year

"Globalizing Modeling Languages" appears in IEEE Computer Magazine. This paper synthesizes our vision of how domain-specific languages form the foundations of global software development. Its appearance in a highly visible venue is major milestone for the dissemination and impact of our work about the diversity of languages.

DiverSE extremely present at the SPLC conference. SPLC is the main international conference for software product line engineering. In 2014, the DiverSE team had a very strong presence at this conference, presenting novel scientific contributions, results of industrial collaborations, and demonstrations of latest software tools.

6.2. Results on Software Language Engineering

The engineering of systems involves many different stakeholders, each with their own domain of expertise. Hence more and more organizations are adopting Domain Specific Languages (DSLs) to allow domain experts to express solutions directly in terms of relevant domain concepts. This new trend raises new challenges about designing DSLs, evolving a set of DSLs and coordinating the use of multiple DSLs for both DSL designers and DSL users. In [56] we present the overall vision that we develop in the DiverSE team about Software Language Engineering. The main results on this topic are presented below.

6.2.1. Globalization of Domain Specific Languages

In the software and systems modeling community, research on domain-specific modeling languages (DSMLs) focuses on technologies for developing languages and tools to increase the effectiveness of domain experts. Yet, there is a lack of support to explicitly relate concepts expressed in different DSMLs, which prevents software and system engineers to reason about information spread across models describing different system aspects. Supporting coordinated use of DSMLs leads to what we call the globalization of modeling languages [20]. In such a context, we develop a research initiative that broadens the DSML research focus beyond the development of independent DSMLs to one that supports globalized DSMLs, that is, DSMLs that facilitate coordination of work across different domains of expertise. We also provid a formal framework to prove the correctness of model driven engineering composition operators [57].

6.2.2. Meta-Language for the Concurrency Concern in DSLs

Concurrency is of primary interest in the development of complex software-intensive systems, as well as the deployment on modern platforms. However, reifying the definition of the DSL concurrency remains a challenge. This hinders: a) the development of a complete understanding of the DSL semantics; b) the effectiveness of concurrency-aware analysis techniques; c) the analysis of the deployment on parallel architectures. In this context, we present MoCCML, a dedicated meta-language for formally specifying the concurrency concern within the definition of a DSL [44]. The concurrency constraints can reflect the knowledge in a particular domain, but also the constraints of a particular platform. MoCCML comes with a complete language workbench to help a DSL designer in the definition of the concurrency directly within the concepts of theDSL itself, and a generic workbench to simulate and analyze any model conforming to this DSL. Mo CCML is illustrated on the definition of an lightweight extension of SDF (SynchronousData Flow).

6.2.3. Automating Variability Model Inference for Component-Based Language Implementations

Componentized language frameworks, coupled with variability modeling, have the potential to bring language development to the masses, by simplifying the configuration of a new language from an existing set of reusable components. However, designing variability models for this purpose requires not only a good understanding of these frameworks and the way components interact, but also an adequate familiarity with the problem domain. In [68] we propose an approach to automatically infer a relevant variability model from a collection of already implemented language components, given a structured, but general representation of the domain. We describe techniques to assist users in achieving a better understanding of the relationships between language components, and find out which languages can be derived from them with respect to the given domain.

6.2.4. Metamorphic Domain-Specific Languages

External or internal domain-specific languages (DSLs) or (fluent) APIs? Whoever you are – a developer or a user of a DSL – you usually have to choose side; you should not! What about metamorphic DSLs that change their shape according to your needs? Our 4-years journey of providing the "right" support (in the domain of feature modeling), led us to develop an external DSL, different shapes of an internal API, and maintain all these languages. A key insight is that there is no one-size-fits-all solution or no clear superiority of a solution compared to another. On the contrary, we found that it does make sense to continue the maintenance of an external and internal DSL. Based on our experience and on an analysis of the DSL engineering field, the vision that we foresee for the future of software languages is their ability to be self-adaptable to the most appropriate shape (including the corresponding integrated development environment) according to a particular usage or task. We call metamorphic DSL such a language, able to change from one shape to another shape [27].

6.2.5. Adapting mutation testing for model transformations

Due to the specificities of models and transformations, classical software testing techniques have to be adapted. Among these techniques, mutation analysis has been ported and a set of mutation operators has been defined. However, mutation analysis currently requires a considerable manual work and is hampered by the test data set improvement activity. This activity is seen by testers as a difficult and time-consuming job, and reduces the benefits of the mutation analysis.

We provide a model transformation traceability mechanism, in conjunction with a model of mutation operators and a dedicated algorithm, to automatically or semi-automatically produce test models that detect new faults [18].

6.2.6. Efficient model cloning for analysis

We propose an original approach that exploits the fact that operations rarely modify a whole model. Given a set of immutable properties, our cloning approach determines the objects and fields that can be shared between the runtime representations of a model and its clones. Our generic cloning algorithm is parameterized with three strategies that establish a trade-off between memory savings and the ease of clone manipulation. We evaluated memory footprints and computation overheads with 100 randomly generated metamodels and models [40]. We have also drawn the research roadmap to exploit these efficient clone operations to analyze multidimensional execution traces [41].

6.3. Results on Variability Modeling and Engineering

6.3.1. Engineering Interactive Systems

In agreement with our permanent effort to validate the techniques we propose on real use cases in various domains, we applied seminal MDE to interactive systems engineering. This led to two collaborations. The first one has been conducted with 3D Collaborative Virtual Environments (3D CVE) researchers. Despite the increasing use of 3D CVE, their development is still a cumbersome task. The various concerns to consider (distributed system, 3D graphics, *etc.*) complexify their development as well as their evolution. We propose to leverage MDE for developing 3D CVEs [45]. We have shown how a 3D CVE framework benefits from a DSL we built using state-of-the-art MDE technologies. The benefits are multiple: 3D CVEs designers can focus on the behavior of their virtual objects without bothering with distributed and graphics features; configuring the content of 3D CVEs and their deployment on various software and hardware platforms can be automated through code generation.

The second collaboration is international and has been conducted with software visualization researchers. Current metamodel editing tools are based on standard visualization and navigation features, such as physical zooms. However, as soon as metamodels become larger, navigating through large metamodels becomes a tedious task that hinders their understanding. In this work, we promote the use of model slicing techniques [102] to build visualization techniques dedicated to metamodels [37]. This approach is implemented in a metamodel visualizer, called `Explain`.

6.3.2. Variability management in regulatory requirements and system engineering

Nuclear power plants are some of the most sophisticated and complex energy systems ever designed. These systems perform safety critical functions and must conform to national safety institutions and international regulations. In many cases, regulatory documents provide very high level and ambiguous requirements that leave a wide margin for interpretation. As the French nuclear industry is now seeking to spread its activities outside France, it is but necessary to master the ins and the outs of the variability between countries safety culture and regulations. This sets both an industrial and a scientific challenge to introduce and propose a product line engineering approach to an unaware industry whose safety culture is made of interpretations, specificities, and exceptions. We have developed two contributions within the French R&D project CONNEXION, while introducing variability modeling to the French nuclear industry [66], [34].

As part of the VaryMDE project (a bilateral collaboration between Thales and Inria) we have developed techniques to generate counter-examples (also called anti-patterns) of model-based product lines [22]. The goal is to infer (1) guidelines or domain-specific rules to avoid earlier the specification of incorrect mappings (2) testing oracles for increasing the robustness of derivation engines given a modeling language. We have applied the approach in the context of a real industrial scenario with Thales involving a large-scale metamodel.

6.3.3. Handling testing challenges in product line engineering

Testing techniques in industry are not yet adapted for product line engineering (PLE).

We have developed original contributions to adapt model-based testing for PLE [65], [63], [13]. We equip usage models, a widely used formalism in MBT, with variability capabilities. Formal correspondences are established between a variability model, a set of functional requirements, and a usage model. An algorithm then exploits the traceability links to automatically derive a usage model variant from a desired set of selected features. The approach is integrated into the MBT tool MaTeLo and is currently used in industry.

We have also developed a variability-based testing approach to derive video sequence variants. The ideas of our VANE approach are i) to encode in a variability model what can vary within a video sequence; ii) to exploit the variability model to generate testable configurations; iii) to synthesize variants of video sequences corresponding to configurations. VANE computes T-wise covering sets while optimizing a function over attributes [50], [25].

6.3.4. Reverse engineering variability models

We have developed automated techniques and a comprehensive environment for synthesizing feature models from various kinds of artefacts (e.g. propositional formula, dependency graph, FMs or product comparison matrices). Specifically we have elaborated a support (through ranking lists, clusters, and logical heuristics) for choosing a sound and meaningful hierarchy [42]. We have performed an empirical evaluation on hundreds of feature models, coming from the SPLOT repository and Wikipedia [108]. We have showed that a hybrid approach mixing logical and ontological techniques outperforms state-of-the-art solutions (to appear in Empirical Software Engineering journal in 2015 [19]). Beyond the reverse engineering of variability, our work has numerous practical applications (e.g., merging multiple product lines, slicing a configuration process).

6.3.5. Product comparison matrices

Product Comparison Matrices (PCMs) constitute a rich source of data for comparing a set of related and competing products over numerous features. Despite their apparent simplicity, PCMs contain heterogeneous, ambiguous, uncontrolled and partial information that hinders their efficient exploitations. We have first elaborated our vision and identify research challenges for an exploitation of PCMs when engineering comparators, configurators, or other services [67].

We have formalized PCMs through model-based automated techniques and developed additional tooling to support the edition and re-engineering of PCMs [43]. 20 participants used our editor to evaluate our PCM metamodel and automated transformations. The empirical results over 75 PCMs from Wikipedia show that (1) a significant proportion of the formalization of PCMs can be automated: 93.11% of the 30061 cells are correctly formalized; (2) the rest of the formalization can be realized by using the editor and mapping cells to existing concepts of the metamodel.

The ASE'2014 paper opens avenues for engaging a community in the mining, re-engineering, edition, and exploitation of PCMs that now abound on the Internet. We have launched an open, collaborative initiative towards this direction http://www.opencompare.org

6.4. Results on Heterogeneous and dynamic software architectures

This year, we focused on the challenges that use *models@runtime* for resource-constrained and resource-aware systems. Our main results are in the following four subdomains:

- We designed an adaptive monitoring framework for component-based systems in which we highlight the benefits of using models@runtime for adaptive monitoring.
- We improved models@runtime technologies for resource-constrained devices.
- We designed efficient reasoning techniques for dynamic software architecture, focusing in particular on resource consumption optimization challenges.
- We performed several experiments on the Internet of Things application domain.

The next section details our experiments.

6.4.1. Resource-aware dynamic architecture

Modern component frameworks support continuous deployment and simultaneous execution of multiple software components on top of the same virtual machine. However, isolation between the various components is limited. A faulty version of any one of the software components can compromise the whole system by consuming all available resources. We propose a solution to efficiently identify faulty software components running simultaneously in a single virtual machine. It is based on an optimistic adaptive monitoring system to identify the faulty component. Suspected components are instrumented to obtain fined grain data for deeper analysis by the monitoring system, but only when required. Unsuspected components are left untouched and execute normally. Thus, we perform localized, just-in-time monitoring that decreases the accumulated overhead of the monitoring system. We evaluated our approach against a state-of-the-art monitoring system and we have shown that our technique correctly detects faulty components, while reducing overhead by an average of 80% [52]. Based on this work, we have presented two tutorials at the CBSE/QoSA conference [49] and at the Middleware conference [51].

6.4.2. Technology enablers for resource-aware dynamic software architecture

Models@runtime provides semantically rich reflection layers enabling intelligent systems to reason about themselves and their surrounding context. Most reasoning processes require not only to explore the current state, but also the past history to take sustainable decisions e.g. to avoid oscillating between states. Models@runtime and model-driven engineering in general lack native mechanisms to efficiently support the notion of history, and current approaches usually generate redundant data when versioning models, which reasoners need to navigate. Because of this limitation, models fail in providing suitable and sustainable abstractions to deal with domains relying on history-aware reasoning. This work tackles this challenge by considering history as a native concept for modeling foundations. Integrated in conjunction with lazy load/storage techniques into the Kevoree Modeling Framework, we demonstrated onto a energy-aware smart grid case study that this mechanisms enable a sustainable reasoning about massive historized models [53].

In this field we also created a specific extension to the `docker.io` open-source project to support a dynamic resource reservation of running containers [9]

6.4.3. Efficient reasoning techniques for dynamic software architecture

Providing software with the capacity of adapting itself according to its environment requires effective techniques to reason and decide on what adaptation to undertake over the running system. To decide on a system adaptation, we have to characterize the value of the system in its corresponding execution environment. A system cannot be characterized by a single dimension, but only using several dimensions such as performance, energy consumption, security and so on. In this context, we have proposed various techniques to leverage multi-objective evolutionary algorithms both at deployment time [46], [21] and at runtime [47] to enable system optimization using multidimensional optimization. We have also proposed a technique to adapt a system proactively based on predictions in order to prevent failures [60]

6.4.4. The Internet of Things application domain

We apply our techniques for heterogeneous and dynamic software architecture more specifically to the Internet of Things application domain. We have two main contributions: (1) an application of the models@runtime concepts on embedded nodes with very limited resources for memory, CPU and battery [30], and (2) a study on the problem of renewable energy production and consumption at home [39]. Domestic microgeneration is the onsite generation of low and zero-carbon heat and electricity by private households to meet their own needs. In this paper we explore how an everyday household routine (doing laundry) can be augmented by digital technologies to help households with photovoltaic solar energy generation to make better use of self-generated energy. We present an 8 month in the field study that involved 18 UK households in longitudinal energy data collection, prototype deployment and participatory data analysis [38]. Through a series of technology interventions mixing energy feedback, proactive suggestions and direct control, the study uncovered opportunities, potential rewards and barriers for families to shift energy consuming household activities. The study highlights how digital technology can act as a mediator between household laundry routines and energy demand-shifting behaviors. Finally, the study provides insights into how a “smart” energy-aware washing machine shapes organization of domestic life and how people “communicate” with their washing machine.

6.5. Results on Diverse Implementations for Resilience

Diversity is acknowledged as a crucial element for resilience, sustainability and increased wealth in many domains such as sociology, economy and ecology. Yet, despite the large body of theoretical and experimental science that emphasizes the need to conserve high levels of diversity in complex systems, the limited amount of diversity in software-intensive systems is a major issue. This is particularly critical as these systems integrate multiple concerns, are connected to the physical world through multiple sensors, run eternally and are open to other services and to users. Here we present our latest observational and technical results about new approaches to increase diversity in software systems.

6.5.1. Automatic synthesis of computationally diverse program variants

The predictability of program execution provides attackers with a rich source of knowledge that they can exploit to spy or remotely control the program. Moving target defense addresses this issue by constantly switching between many diverse variants of a program, thus reducing the certainty that an attacker can have about the program execution. The effectiveness of this approach relies on the availability of a large number of software variants that exhibit different executions. However, current approaches rely on the natural diversity provided by off-the-shelf components, which is very limited. We have explored the automatic synthesis of large sets of program variants, called *sosies* [32]. Sosies provide the same expected functionality as the original program, while exhibiting different executions. They are said to be computationally diverse.

6.5.2. Software Evolution for Diversity Emergence

We aim at favoring spontaneous diversification in software systems, to increase their adaptive capacities. This objective is founded on three observations: (1) software has to constantly evolve to face unpredictable changes in its requirements, execution environment or to respond to failure (bugs, attacks, etc.); (2) the emergence and maintenance of high levels of diversity are essential to provide adaptive capacities to many forms of complex systems, ranging from ecological and biological systems to social and economical systems; (3) diversity levels tend to be very low in software systems. In this work [33], we consider evolution as a driver for diversity as a means to increase resilience in software systems. In particular, we are inspired by bipartite ecological relationships to investigate the automatic diversification of the server side of a client-server architecture.

6.5.3. Analyzing the diversity of development practices in open source projects

Decentralized version control systems allow a rich structure of commit histories, which presents features that are typical of complex graph models. We bring some evidences of how the very structure of these commit histories carries relevant information about the distributed development process. By means of a novel data structure that we formally define, we analyze the topological characteristics of commit graphs of a sample of git projects. Our findings point out the existence of common recurrent structural patterns that identically occur in different projects and can be considered building blocks of distributed collaborative development [36], [35].

DOLPHIN Project-Team

6. New Results

6.1. Highlights of the Year

In [23], we have revisited the design and implementation of the Branch and Bound algorithm for solving on large scale distributed environments challenging permutation-based optimization problems such as Q3AP. The new approach includes original ways to efficiently deal with some crucial issues mainly, dynamic adaptive load balancing and fault tolerance. The approach allowed to solve to optimality for the first time a difficult Q3AP instance (Nug15) on the nation-wide Grid'5000 computational grid. The resolution was completed within less than 12 days using an average of 1,123 processing cores distributed over 6 Grid'5000 sites and peaked at 3,427.

6.2. Fitness Landscape Analysis for Multiobjective Optimization

Participant: Arnaud Liefooghe.

The properties of local optimal solutions in multi-objective combinatorial optimization problems are crucial for the effectiveness of local search algorithms, particularly when these algorithms are based on Pareto dominance. Such local search algorithms typically return a set of mutually non-dominated Pareto local optimal (PLO) solutions, that is, a PLO-set. In [34], we investigate two aspects of PLO-sets by means of experiments with Pareto local search (PLS). First, we examine the impact of several problem characteristics on the properties of PLO-sets for multi-objective NK-landscapes with correlated objectives. In particular, we report that either increasing the number of objectives or decreasing the correlation between objectives leads to an exponential increment on the size of PLO-sets, whereas the variable correlation has only a minor effect. Second, we study the running time and the quality reached when using bounding archiving methods to limit the size of the archive handled by PLS, and thus, the maximum size of the PLO-set found. We argue that there is a clear relationship between the running time of PLS and the difficulty of a problem instance.

Complementarily, in [25] we study the behavior of three elitist multi- and many-objective evolutionary algorithms in generating a high-resolution approximation of the Pareto set. Several search-assessment indicators are defined to trace the dynamics of survival selection and measure the ability to simultaneously keep optimal solutions and discover new ones under different population sizes, set as a fraction of the Pareto set size. Our study clarifies the ability and efficiency of the algorithms assuming scenarios where it is relatively easy to hit the Pareto set, showing the importance to properly assess algorithm's performance according to the task of the optimizer in many-objective optimization.

6.3. Combining dynamic programming and metaheuristics for the Unit Commitment Problem

Participants: Sophie Jacquin, Laetitia Jourdan, El-Ghazali Talbi.

DYNAMOP (DYNAmic programming using Metaheuristic for Optimization Problems) is a new dynamic programming based on genetic algorithm. It uses a representation based on a path in the graph of states of dynamic programming which is adapted to dynamic structure of the problem and facilitates the hybridization between evolutionary algorithms and dynamic programming. Experiments indicate that the proposed approach outperforms the best known in literature [44].

6.4. Multi-decoding strategy for Multi-objective Unit Commitment Problem

Participants: Sophie Jacquin, Lucien Mousin, Igor Machado, Laetitia Jourdan, El-Ghazali Talbi.

In the multiobjective version of the UCP taking the emission of gas into account, the dispatching problem remains easy to solve whereas considering it separatly remains interesting. A multi-objective GA handling binary vectors is applied. However for a binary representation there is a set of solutions of the dispatching problem that are pareto equivalent. In this approach a genotypic solution is associated with a set of phenotypic solutions. This set of solutions is from the optimal pareto front solution of the dispatching problem associated with the genotypic solution. As many phenotypic solutions are attached to a single genotypic solution, the fitness assignment and diversity assignment methods of NSGA-II have to be adapted. The multi decoding embedded approach has shown very good performances in comparison to two other less complex decoding systems.

6.5. Decomposition-Based Algorithms for Multiobjective Optimization

Participants: Dimo Brockhoff, Bilel Derbel, Arnaud Liefooghe, Gauvain Marquet, El-Ghazali Talbi.

Recently, there has been a renewed interest in decomposition-based approaches for evolutionary multiobjective optimization. Those algorithms decompose a multiobjective optimization problem into several single-objective optimization problems by using so-called scalarizing functions which are then simultaneously optimized by single-objective algorithms in a cooperative manner.

Our contributions to decomposition-based algorithms in 2014 has been three-fold. Firstly, we investigated in [28] the general impact of different scalarizing functions and their parameters on the search performance. We thereby abstracted from any specific algorithm and only considered the difficulty of the single scalarized problems in terms of the search ability of a (1+lambda)-EA on bi-objective NK-landscapes. Secondly, in [16], we proposed a new distributed heuristic for approximating the Pareto set of bi-objective optimization problems. Given a number of computing nodes, we self-coordinate them locally, in order to cooperatively search different regions of the Pareto front. As local information, every node uses only the positions of its neighbors in the objective space and evolves its local solution adaptively, based on what we term a ‘localized fitness function’. We deployed our distributed algorithm using a computer cluster of hundreds of cores. At last, we enhanced the algorithm MOEA/D, a prominent example of a decomposition-based algorithm from the literature, by investigating the idea of evolving the whole population simultaneously at once. We thereby proposed new alternative selection and replacement strategies that can be combined in different ways within a generic and problem-independent framework [36].

6.6. Link-Heterogeneous work stealing for Branch-and-Bound Algorithms

Participants: T-T Vu, Bilel Derbel.

In this work [41], we push forward the design of parallel and distributed optimization algorithms running on link-heterogeneous systems where network latencies can deeply impact performance. We consider parallel Branch-and-Bound (B&B), viewed as a generic algorithm searching in a dynamic tree representing a set of candidate solutions built dynamically. A major challenge is then to deal with the irregularity of B&B computations and to distribute workload evenly at runtime. In this context, the random work-stealing paradigm has been proved to be extremely beneficial. However, it is known to perform loosely in non-homogeneous distributed systems where communications costs are a major obstacle for high performance. We there-by investigate the design of an effective work-stealing protocol dealing with the heterogeneity of network link latencies. We propose a generic distributed algorithm which can be easily implemented to fit different types of heterogeneity. The proposed algorithm extends on reference approaches, namely Probabilistic Work Stealing (PWS), and Adaptive Cluster-aware Random Stealing (ACRS); by introducing new adaptive control operations that are shown to be highly accurate in increasing work locality and decreasing steals cost. Through emulations on top of a real test-bed, we provide a comprehensive experimental analysis including: (i) a comparative study on a broad range of harsh network scenarios going from flat networks to more hierarchical grid-like networks, and (ii) an in-depth analysis of protocols’ behavior at the aim of gaining new insights into dynamic load-balancing in heterogeneous distributed environments. Over all experimented configurations, our results show that although the proposed protocol is not tailored for a specific networked platform, it can save 30% execution time in average compared to its competitors, while demonstrating high quality self-adjusting capabilities.

6.7. New data structure for solving large permutation problems using multi-core B&B

Participants: Rudi Leroy, Nouredine Melab.

Solving large permutation problems using parallel B&B algorithms results in the generation of a very large pool of subproblems. Defining an efficient data structure is highly required to store and manage efficiently that pool. In [31], we have proposed a new dedicated data structure called *Integer-Vector-Matrix* (or *IVM* and redefined the operators of the B&B algorithm acting on it. We have also revisited the Work Stealing mechanism on multi-core processors. In the proposed approach, work units are coded in a coalesced way using factoradic-based intervals, and private IVMs are used to store and explore locally subsets of subproblems. The IVM-based approach has been experimented and compared to the approach based on concurrent linked-list, which is often used. The results show that our approach is more efficient in terms of memory usage and management time. In [31], we have investigated various work stealing strategies based on different victim selection and granularity policies. This later paper has been selected for a special issue in the CCPE international journal.

6.8. B&BGrid revisited for solving challenging Q3AP instances on large volatile computational environments

Participants: Nouredine Melab, El-Ghazali Talbi.

We have revisited the design and implementation of parallel B&B algorithms on multi-core (collaboration with UMONS, Belgium) and grid-wide environments (collaboration with University of Luxembourg and UMONS, Belgium)) for solving to otimality and efficiently large permutation problem instances. We have proposed a gridification approach of the B&B algorithm called B&B@Grid. This later includes a dynamic load balancing technique and a checkpointing mechanism for permutation problems. The approach has been validated through single-permutation Flowshop problem. In [23], we have extended the approach to deal with more than one permutation. To do that, we have revisited the design and implementation of the dynamic load balancing and checkpointing mechanisms for multiple permutation-problems. The new approach allowed the optimal resolution on a nation-wide grid (Grid'5000) of a difficult instance of the 3D quadratic assignment problem (Q3AP). To solve the instance, an average of 1,123 processing cores were used during less than 12 days with a peak of around 3,427 CPU cores.

DRACULA Project-Team

6. New Results

6.1. Highlights of the Year

- Marine Jacquier and Fabien Crauste (in collaboration with C.O. Soulage and H.A. Soul) published a paper ([18], see also § 6.7) in PLoS ONE 2014.
- Sotiris Prokopiou, Loic Barbarroux, Samuel Bernard, Olivier Gandrillon and Fabien Crauste (in collaboration with J. Mafille, Y. Leverrier, C. Arpin and J. Marvel) published a paper ([21], see also § 6.2) in Computation 2014.
- We organized a session "Deterministic and stochastic models in biology and medicine" at 10th AIMS Conference on Dynamical Systems, Differential Equations and Applications, Madrid (Spain), 7 - 11 July 2014 http://www.aimsciences.org/conferences/2014/.
- Our project entitled "Prion and Alzheimer: mathematical modeling and experiments dealing with a dangerous liaison" has been granted by the French Association France Alzheimer, and has been selected with 3 other projects amongst 14 supported works to be part of a scientific popularizing broadcasting campaign through a short scientific cartoon http://www.francealzheimer.org/projets-soutenus-cette-ann%C3%A9e/lab-alz-comprendre-enjeux-recherche/964 and https://www.youtube.com/watch?v=X0mLf8IJhV4&list=PLCq-e7n2r6Wgo3kaseDHetNAPAG7y9B-d.

6.2. Multi-scale model of the CD8 T cell immune response

We presented in [21] the first multi-scale model of CD8 T cell activation in a lymph node, following an acute infection. CD8 T cell dynamics are described using a cellular Potts model (hence cells are discrete interacting objects), whereas intracellular regulation is associated with a continuous system of nonlinear ordinary differential equations focusing on the dynamics of key proteins. This model allows to reproduce the dynamics of CD8 T cells over a five days period (corresponding to the activation and differentiation into effector cells) and is currently used to characterize the generation of memory cells.

6.3. Mathematical model of hematopoiesis

We investigate in [5] a mathematical model of blood cell production in the bone marrow (hematopoiesis). The model describes both the evolution of primitive hematopoietic stem cells and the maturation of these cells as they differentiate to form the three kinds of progenitors and mature blood cells (red blood cells, white cells and platelets). The three types of progenitors and mature cells are coupled to each other via their common origin in primitive hematopoietic stem cells compartment. The resulting system is composed by eleven age-structured partial differential equations. To analyze this model, we don't take into account cell age-dependence of coefficients, that prevents a usual reduction of the structured system to an unstructured delay differential system. We study the existence of stationary solutions: trivial, axial and positive steady states. Then we give conditions for the local asymptotic stability of the trivial steady state and by using a Lyapunov function, we obtain a sufficient condition for its global asymptotic stability. In some particular cases, we analyze the local asymptotic stability of the positive steady state by using the characteristic equation. Finally, by numerical simulations, we illustrate our results and we show that a change in the duration of cell cycle can cause oscillations.

6.4. The role of spatial organization of cells in erythropoiesis

Erythropoiesis, the process of red blood cell production occurs mainly in the bone marrow. The functional unit of mammalian erythropoiesis, the erythroblastic island, consists of a central macrophage surrounded by adherent erythroid progenitor cells (CFU-E/Pro-EBs) and their differentiating progeny, the erythroblasts. Central macrophages display on their surface or secrete various growth or inhibitory factors that influence the fate of the surrounding erythroid cells. CFU-E/Pro-EBs have three possible fates : a) expansion of their numbers without differentiation, b) differentiation into reticulocytes that are released into the blood, c) death by apoptosis. CFU-E/Pro-EB fate is under the control of a complex molecular network, that is highly dependent upon environmental conditions in the erythroblastic island. In order to assess the functional role of space coupled with the complex network behavior in erythroblastic islands, we developed hybrid discrete-continuous models of erythropoiesis. In [13], a model was developed in which cells are considered as individual physical objects, intracellular regulatory networks are modeled with ordinary differential equations and extracellular concentrations by partial differential equations. We used this model to investigate the impact of an important difference between humans and mice in which mature late-stage erythroblasts produce the most Fas-ligand in humans, whereas early-stage erythroblasts produce the most Fas-ligand in mice. Although the global behaviors of the erythroblastic islands in both species were similar, differences were found, including a relatively slower response time to acute anemia in humans. Also, our modeling approach was very consistent with in vitro culture data, where the central macrophage in reconstituted erythroblastic islands has a strong impact on the dynamics of red blood cell production. Conclusions: The specific spatial organization of erythroblastic islands is key to the normal, stable functioning of mammalian erythropoiesis, both in vitro and in vivo. Our model of a simplified molecular network controlling cell decision provides a realistic functional unit of mammalian erythropoiesis that integrates multiple microenvironmental influences within the erythroblastic island with those of circulating regulators of erythropoiesis, such as EPO and glucocorticosteroids, that are produced at remote sites.

6.5. Mathematical modelling of cell polarization

In [19], a fine description of the behaviour of a nonlinear drift diffusion model inspired from spontaneous cell polarization was performed. This model has Keller Segel type properties and in particular, quantitative proofs were obtained for the convergence to steady state or self similar profile or blow up. The behaviour depends on the mass of the initial data.

6.6. Numerical modelling of cell distribution in blood flow

Properties of blood cells and their interaction determine their distribution in flow. It is observed experimentally that erythrocytes migrate to the ow axis, platelets to the vessel wall, and leucocytes roll along the vessel wall. In [2], a three-dimensional model based on Dissipative Particle Dynamics method and a new hybrid (discrete-continuous) model for blood cells is used to study the interaction of erythrocytes with platelets and leucocytes in flow. Erythrocytes are modelled as elastic highly deformable membranes, while platelets and leucocytes as elastic membranes with their shape close to a sphere. Separation of erythrocytes and platelets in flow is shown for different values of hematocrit. Erythrocyte and platelet distributions are in a good qualitative agreement with the existing experimental results. Migration of leucocyte to the vessel wall and its rolling along the wall is observed.

6.7. Mathematical model of food intake dynamics

In [18], we propose a nonlinear mathematical model of food intake dynamics and body weight dynamics, involving the description of several regulating hormones (leptin, ghrelin, insulin). Using a temporal perturbation of food availability in groups of rats, this model is able to predict body weight and food intake variations by taking into account energy expenditure dynamics based on a memory of the previous food intake. This model also allowed us to estimate the memory lag to approximately 8 days. It also explains how important variations in food availability during periods longer than these 8 days can induce body weight gains.

6.8. Long time existence of weak solutions to cross diffusion models

We pointed out a general entropy structure in cross diffusion systems. We used this structure with duality arguments to build a general framework in which weak solutions exist for a long time. This led to two research articles (one [10] in Siam Journal of Mathematical Analysis and one and [32] recently accepted in Comm. In PDE). This was conducted with the help of the ANR KIBORD.

6.9. Mathematics of Darwin's diagram

Darwin illustrated his theory about emergence and evolution of biological species with a diagram. It shows how species exist, evolve, appear and disappear. Our goal in [8] is to give a mathematical interpretation of this diagram and to show how it can be reproduced in mathematical models. It appears that conventional models in population dynamics are not sufficient, and we introduce a number of new models which take into account local, nonlocal and global consumption of resources, and models with space and time dependent coefficients.

6.10. A micellar on-pathway intermediate step explains the kinetics of prion amyloid formation

In [16], we used a strong interdisciplinary collaboration between mathematicians and biologists to exhibit a new element taking an important role in the development of the pathological prion formation. Indeed, in a previous work by Alvarez-Martinez et al. (2011), the authors pointed out some fallacies in the mainstream interpretation of the prion amyloid formation. It appeared necessary to propose an original hypothesis able to reconcile the in vitro data with the predictions of a mathematical model describing the problem. Here, a model is developed accordingly with the hypothesis that an intermediate on-pathway leads to the conformation of the prion protein into an amyloid competent isoform thanks to a structure, called micelles, formed from hydrodynamic interaction. The authors also compared data to the prediction of their model and proposed a new hypothesis for the formation of infectious prion amyloids.

DREAM Project-Team

6. New Results

6.1. Simulator-based decision support

Participants: Louis Bonneau de Beaufort, Tassadit Bouadi, Marie-Odile Cordier, Thomas Guyet, Christine Largouët, Véronique Masson, René Quiniou, Sophie Robin, Laurence Rozé, Yulong Zhao.

6.1.1. Model-checking an ecosystem for decision-aid

In previous work we have proposed to use qualitative modelling to model ecosystems and we defined a set of high level query patterns to explore th system [53]. This approach has been applied on real-case ecosystems (coral-reef ecosystem in New-Caledonia, fisheries ecosystem in the English channel) and implemented in a tool called EcoMata.

In recent studies we have focussed on the formalization of the qualitative model automatically built from an abstracted ecosystem description. Ecosystems share some common features with concurrent systems represented in the model-checking field: the system complexity is due to interacting components and the system evolution is event-driven and submitted to temporal constraints. However if model-checking techniques are dedicated to finite state systems, ecosystem models are usually represented by analytical models as a set of differential equations. Some studies present how to quantize continuous-time systems in order to diagnose them as discrete-event systems. We proposed a method to build automatically a network of timed automata from various information on the system: description of interactions between components, human knowledge, simple models of population dynamics. The key point is to quantize the continuous-time sub-systems and to get a qualitative model described as network of timed automata. To reduce the size of this network, important after the automatic generation, a learning machine algorithm has been applied in order to reduce the number of "similar" locations. This work has been published in [37].

6.1.2. Controler synthesis for optimal strategy search

Similarly to previous work, this approach relies on a qualitative model of a dynamical system. The problem consists in finding a strategy in order to help the user achieveing a specific goal. The model is now considered as a timed game automata expressing controllable and uncontrollable actions. The strategy represents the sequence of actions that can be performed by a user to reach a particular state (in case of a reachability problem for instance). A first approach based on a "generate and test" method has been developped for the marine ecosystem example [69].

More recently, two new methods for finding the optimal strategies have been proposed. The first one uses controller synthesis on timed automata and exploits the efficency of well-recognized tools. The second one deals with a set of similar models and extracts a more general strategy, closer to what is expected by the stakeholders. These methods have been applied in the context of herd management on a catchment. Yulong Zhao defended his Phd this year on this research subject [5].

6.1.3. A datawarehouse for simulation data

In previous work we have proposed a datawarehouse architecture to store the huge data produced by deep agricultural simulation models [35]. This year, we have worked on hierarchical skyline queries to introduce skyline queries in a datawarehouse framework. Conventional skyline queries retrieve the skyline points in a context of dimensions with a single hierarchical level. However, in some applications with multidimensional and hierarchical data structure (e.g. data warehouses), skyline points may be associated with dimensions having multiple hierarchical levels. Thus, we have proposed an efficient approach reproducing the effect of the OLAP operators "drill-down" and "roll-up" on the computation of skyline queries [10], [25]. It provides the user with navigation operators along the dimensions hierarchies (i.e. specialize / generalize) while ensuring an online calculation of the associated skyline.

6.1.4. Post-mining classification rules

We consider sets of classification rules with quantitative attributes inferred by supervised machine learning, as in the framework of the SACADEAU project. Our aim is to improve human understanding of such sets of rules. Often, output quantitative rules contains too many intervals that are difficult to intepret. It is thus important to merge some of these intervals in order to get more understandable rules. However, blindly merging rules may decrease rule quality. To counter that, we proposed two algorithms for merging intervals via clustering techniques that take into account the final rule quality. The approach automatically detects the most adapted number of clusters required to merge intervals while maintaining rule quality.

6.2. Data Mining

Participants: Sid Ahmed Benabderrahmane, Marie-Odile Cordier, Serge Vladimir Emteu Tchagou, Thomas Guyet, Yves Moinard, René Quiniou, Alexandre Termier.

6.2.1. Application of sequential pattern mining with intervals

Our theoretical work on sequential pattern mining with intervals [47] has been applied to two real issues: the customer relationship management and analysis of care pathways.

Customer Relationship Management (CRM) comprises a set of tools for managing the interactions between a company and its customers. The main objective of the data analysts is to propose the correct service to a customer at the correct moment by applying decision rules. If rules or sequential patterns can predict the interaction that can follow a sequence of actions or events, they can not predict at what time such actions have the highest probability to occur. The objective of temporal pattern mining is to refine the prediction by extracting patterns with information about the duration and delay between the events. This year we have experimented two algorithms on a CRM databases, QTIPrefixSpan [47] and TGSP [68], to extract sequential patterns with quantitative temporal information. We have integrated the TGSP algorithm into an interface to visualize and to browse the extracted patterns. A paper describing this contribution have been recently accepted in a workshop [41].

The QTIPrefixSpan algorithm has also been applied to the analysis of care pathways. The pharmaco-epidemiology platform of the Rennes hospital was interested in characterizing the care pathways preceeding the epileptic seizures of stable epilepic patients. A care pathway consist of the sequence of drug exposures (temporal intervals). The objective is to study the ability of QTIPrefixSpan to identify drug switches between original and generic anti-epileptic drugs. This work is still in progress and will be extended in the PEPS project (see section 8.1.1).

6.2.2. Multiscale segmentation of satellite image time series

Satellite images allow the acquisition of large-scale ground vegetation. Images are available along several years with a high acquisition frequency (1 image every two weeks). Such data are called satellite image time series (SITS). In [45], we presented a method to segment an image through the characterization of the evolution of a vegetation index (NDVI) on two scales: annual and multi-year. This work is now under submission to the journal on Remote Sensing in Environment. The main issue of this approach was the required computation resources (time and memory). Last year, we applied 1D-SAX to reduce data dimensionality [21]. This approach on the supervised classification of large SITS of Senegal and we showed that 1D-SAX approaches the classification results of time series while significantly reducing the required memory storage of the images.

This year, we first continued to explore the supervised classification of SITS using classification trees for time-series [40] by implementing a parallelized version of this algorithm. Secondly, we explored the adaption of the object-oriented segmentation to SITS. The object-oriented segmentation [34] is able to segment images based on the segment uniformity. We proposed a measure for time-series uniformity to adapt the segmentation algorithm and applied it on large multivariate SITS of Senegal. This work have been presented to the conference on spatial analysis and geography [16]. A collaboration with A. Fall (Université Paris-13) have been initiated to compare our results on the Senegal with ground observations. Moreover, we planned to apply our algorithm to analyse the land use in Peru (collaboration with A. Marshall, Université Paris 13/PRODIG).

6.2.3. Analysis and simulation of landscape based on spatial patterns

Researchers in agro-environment need a great variety of landscapes to test their scientific hypotheses using agro-ecological models. Real landscapes are difficult to acquire and do not enable the agronomist to test all their hypothesis. Working with simulated landscapes is then an alternative to get a sufficient variety of experimental data. Our objective is to develop an original scheme to generate landscapes that reproduce realistic interface properties between parcels. This approach is made of the extraction of spatial patterns from a real geographic area and the use of these patterns to generate new "realistic" landscapes. It is based on a spatial representation of landscapes by a graph expressing the spatial relationships between the agricultural parcels (as well as the roads, the rivers, the buildings, etc.), of a specific geographic area.

In past years, we worked on the exploration of graph mining techniques, such as gSPAN [67], to discover the relevant spatial patterns present in a spatial-graph. We assume that the set of the frequent graph patterns are the characterisation of the landscape. Our remaining challenge was to simulate new realistic landscapes that will reproduce the same patterns.

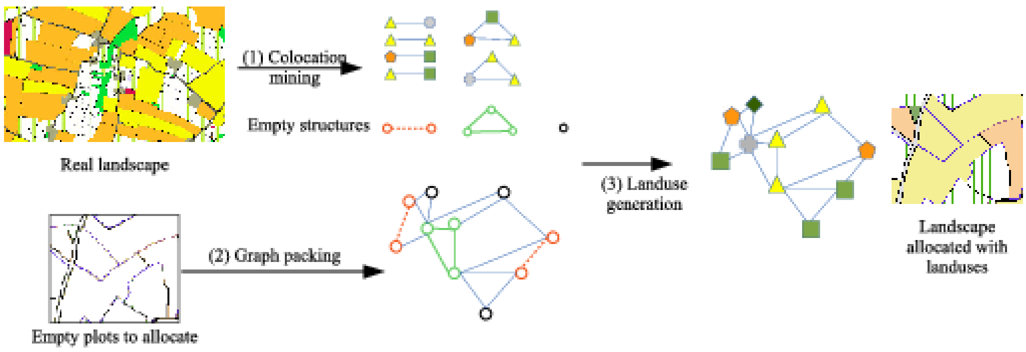


Figure 1. Simulation process in three steps : 1) characteristic graph-patterns mining, 2) graph packing of the cadastral landscape and 3) land use assignement.

This year, we formalized the simulation process by a formal problem of graph packing [51]. The process is illustrated by Figure 1 . Solving instances of the general graph packing problem has a high combinatorics and there does not exists any efficient algorithm to solve it. We proposed an ASP program to tackle the combinatorics of the graph packing and to assign the land use considering some expert knowledge. Our approach combines the efficiency of ASP to solve the packing issue and the simplicity of the declarative programming to take into account the expert contraints on the land use. Contraints about the minimum surface of crops or about the impossibility of some crops colocation can be easily defined. This work have been presented at the conference RFIA [19] and we have been invited to provide an extended version to the Revue d'Intelligence Artificielle (RIA). The application results have been presented to the national colloquium on landscape modelling (http://www.reseau-payote.fr/?q=colloque2014).

In addition to the landscape simulation, the challenging tasks of solving the general graph packing with ASP raises interests in more general problem (such as graph compression). We initiated a collaboration with J. Nicolas (Inria/Dyliss) to improve the efficiency of our first programs.

6.2.4. Mining with ASP

In pattern mining, a pattern is considered interesting if it occurs frequently in the data, i.e. the number of its occurrences is greater than a fixed given threshold. As non informed mining methods tend to generate massive results, there is more and more interest in pattern mining algorithms able to mine data considering some expert knowledge. Though a generic pattern mining tool that could be tailored to the specific task of a data-scientist is still a holy grail for pattern mining software designers, some recent attempts have proposed generic pattern mining tools [44] for itemset mining tasks. In collaboration with Torsten Schaub, we explore the ability of a declarative language, such as Answer Set Programming (ASP), to solve pattern mining tasks efficiently. A first attempt have been proposed by Jarvisälo for simple settings [49].

This year, we worked on several classical pattern mining tasks: episodes, sequences and closed/maximal itemsets. In [20], we explore the use of ASP to extract frequent episodes (without parallel events) in a unique long sequence of itemsets. We especially evaluate the incremental resolution to improve the efficiency of our program. We next worked on sequence mining to extract pattern from the sequence of TV programs (V. Claveau, CNRS/LinkMedia). This tasks was simpler, but the computation time was significantly higher than dedicated algorithms. Nonetheless, our recent programs extracting closed or maximal patterns have better results.

6.2.5. Monitoring cattle

Following the lines of a previous work [62], we are working on a method for detecting Bovine Respiratory Diseases (BRD) from behavioral (walking, lying, feeding and drinking activity) and physiological (rumen temperature) data recorded on feedlot cattle being fattened up in big farms in Alberta (Canada). This year, we have especially worked on multivariate sensor analysis to devising multivariate decision rules for improving the specificity of detectors [15].

6.2.6. Subdimensional clustering for fast similarity search over time series data. Application to Information retrieval tasks

Information retrieval and similarity search tasks in time series databases remains a challenge that requires to discover relevant pattern-sequences that are recurrent over the overall time series sequences, and to find temporal associations among these frequently occurring patterns. However, proposed methods suffer from a lack of flexibility of the used similarity measures, a lack of scalability of the representation model, and a penalizing runtime to retrieve the information. Motivated by these observations, we have designed a framework tackling the query by content problem on time series data, ensuring (i) fast response time, (ii) multi-level information representation, and (iii) representing temporal associations between extracted patterns. This year we have compared several distance measures on time series with different criteria and proposed a hybrid retrieval method based on pattern extraction and clustering [8].

6.2.7. Knowledge Extraction from Heterogeneous Data

Recently, mining microarrays data has become a big challenge due to the growing sources of available data. We are using machine learning methods such as clustering, dimensionality reduction, association rules discovery on transcriptomic data, by combining a domain ontology as source of knowledge, in order to supervise the KDD process. Our objectives concern the identification of genes that could participate in the development of tumors. This year, we have introduced a new method for extracting enriched biological functions from transcriptomic databases using an integrative bi-classication approach based on formal concept analysis [7].

6.2.8. Trace reduction

One problem of execution trace of applications on embedded systems is that they can grow very large, typically several Gigabytes for 5 minutes of audio/video playback. Some endurance tests require continuous playback for 96 hours, which would lead to hundreds of Gigabytes of traces, that current techniques cannot analyze. We have proposed TraceSquiz, an online approach to monitor the trace output during endurance test, in order to record only suspicious portions of the trace and discard regular ones. This approach is based on anomaly detection techniques, and as been accepted in the DATE'15 conference [14].

6.3. Causal reasoning and argumentation

Participants: Philippe Besnard, Marie-Odile Cordier, Yves Moinard.

We have continued our work on reasoning (precisely search for explanations) from causal relations and ontology [9]. Mainly, we have enforced the use or argumentation [32] in order to help choosing the best explanations among the (rather big) set of explanations given by our previous formalism. Then, we hope to be able to use the last versions of clingo in order to get an efficient tool to deal with complex situations (our example is Xynthia storm, february 2012 in western France for which there exists a huge amount of data from various official reports) by using clingo. For now we have a preliminary program which provides (besides the applications already mentioned: mining and landscape simulation) another application of the recent versions of ASP. One interest is that the nature of the rules in ASP should allow to translate rather directly (and hopefully efficiently) our previous formalism together with the improved argumentation part.

DREAMPAL Team

5. New Results

5.1. Highlights of the Year

The papers [4] and [6] are published in journals (Software Testing Verification and Analisys, resp. Formal Aspects of Computing) that are among the best in their respective fields.

5.2. HoMade

HOMADE V5 is available from 03/2014. New features cover :

- new pipeline architecture with delayed conditional branch
- new unified FSM: Pipeline 2 stages
- renumbering of some IPs
- new activity management on the Slaves in 1D / 2D : by the master OnX , OnY, OnXY, and by the slaves the IPsleep removes the Slave from the next SPMDcall
- new bit per bit loading of program memories, for master and slaves
- new names for some components.
- new versions of a lot of IPs (inside)
- new communication network between Slaves: 2D torus ring with broadcast and communication on x or y axis
- new input binary file format (to respect !!)
- new test_bench for fast reading of instruction files
- new UART wrapper
- new assembler Hasm for those that do not speak binary
- nexys3 version for cheap platform experimentation (does not support more than 2x1 Slaves)
- V6 V7 xilinx supports up to 12 x 12 slaves
- Isim supports many more slaves !!!

More details can be found on www.lifl.fr/~dekeyser/Homade.

5.3. HiHope : A higher level language for the HoMade processor

HiHope is a programming language inspired by Forth used to program the HoMade processor. It includes language constructs for switching at runtime between hardware functions (implemented by IPs) and software functions in a transparent way. We also propose the notion of parallel function language construct. As a result, HiHope programs can use either hardware IPs or software functions, and can perform both sequential and parallel function calls, as well as sequential and parallel function redefinitions.

5.4. Integrating Profiling into MDE Compilers

This work [3] aims at improving performance by returning to the high-level models, specific execution data from a profiling tool enhanced by smart advices computed by an analysis engine. In order to keep the link between execution and model, the process is based on a traceability mechanism. Once the model is automatically annotated, it can be re-factored aiming better performances on the re-generated code. Hence, this work allows keeping coherence between model and code without forgetting to harness the power of parallel architectures. The example uses a transformation chain from UML-MARTE models to OpenCL code.

5.5. Language-Independent Symbolic Execution, Program Equivalence, and Program Verification

A significant part of our research project consists in applying formal techniques for symbolically executing and formally verifying HiHope programs, as well as for formally proving the equivalence of HiHope programs with the corresponding HoMade assembly and machine-code programs obtained by compilation of HiHope.

- Symbolic execution will detect bugs (e.g., stack undeflow) in HiHope programs. Additionaly, symbolic execution is the natural execution manner of HiHope programs as soon as they contain (typically, underspecified) hardware IPs;
- program verification will guarantee the absence of bugs (with respect to specified properties, e.g., no stack underflow, no invocation of unavailable IPs, ...);
- program equivalence will guarantee that such above-mentioned bugs are also absent from the HoMade assembly and machine-code programs obtained by compilation of HiHope source code.

Since these languages are still evolving we decided to work (together with our colleagues from Univ. Iasi, Romania) on language-independent symbolic execution, program-equivalence, and program-verification techniques. In this way, when all the languages in our project become stable, we will be readily able to instantiate the above generic techniques on (the K formal definitions of) the languages in question. We note that all the techniques described below are also independent of K: they are applicable to other language-definition frameworks that use similar rewriting-based formal operational semantics.

5.5.1. Symbolic Execution

In [15] we propose a language-independent symbolic execution framework. The approach is parameterised by a language definition, which consists of a signature for the language's syntax and execution infrastructure, a model interpreting the signature, and rewrite rules for the language's operational semantics. Then, symbolic execution amounts to performing a so-called symbolic rewriting, which consists in changing both the model and the manner in which the operational semantics rules are applied. We prove that the symbolic execution thus defined has the properties naturally expected from it. A prototype implementation of our approach was developed in the K Framework. We demonstrate the genericity of our tool by instantiating it on several languages, and show how it can be used for the symbolic execution, bounded model checking, and deductive verification of several programs. With respect to earlier versions of this work, we have redefined symbolic execution in a more generic way and have included applications to model checking and deductive verification. The current version of the report [15] is submitted to a journal and is based on Andrai Arusoaie's PhD thesis [1], defended in September 2014 at Univ. Iasi (Romania). Andrei was co-supervised by Vlad Rusu and has since joined Dreampal as a postdoc.

5.5.2. Program Equivalences

In [6] we propose a logic and a deductive system for stating and automatically proving the equivalence of programs written in languages having a rewriting-based operational semantics. The chosen equivalence is parametric in a so-called observation relation, and it says that two programs satisfying the observation relation will inevitably be, in the future, in the observation relation again. This notion of equivalence generalises several well-known equivalences and is appropriate for deterministic (or, at least, for confluent) programs. The deductive system is circular in nature and is proved sound and weakly complete; together, these results say that, when it terminates, our system correctly solves the given program-equivalence problem. We show that our approach is suitable for proving equivalence for terminating and non-terminating programs as well as for concrete and symbolic programs. The latter are programs in which some statements or expressions are symbolic variables. By proving the equivalence between symbolic programs, one proves the equivalence of (infinitely) many concrete programs obtained by replacing the variables by concrete statements or expressions. The approach is illustrated by proving program equivalence in two languages from different programming paradigms. The examples in the paper, as well as other examples, can be checked using an online tool. This work was started in 2012. With respect to earlier versions, the new journal publication [6] includes a new and more general presentation of program equivalence as a temporal-logic formula, the generalisation of the approach to nondeterministic-confluent language semantics, substantially more compact proofs, and a new application to corecursive programs.

In another work [10] we deal with a different kind of equivalence: *mutual equivalence*, which says that two programs are mutually equivalent if they both diverge or they end up in similar states. Mutual equivalence is an adequate notion of equivalence for programs written in deterministic languages. It is useful in many contexts, such as capturing the correctness of, program transformations within the same language, or capturing the correctness of compilers between two different languages. In the case of different languages one needs an operation called *language aggregation*, which we present in [11] in more detail, that combine two languages into a single one. We introduce a language-independent proof system for mutual equivalence, which is parametric in the operational semantics of two languages and in a state-similarity relation. The proof system is sound: if it terminates then it establishes the mutual equivalence of the programs given to it as input. We illustrate it on two programs in two different languages (an imperative one and a functional one), that both compute the Collatz sequence.

5.5.3. Program Verification

In [16] we present an automatic, language-independent program verification approach and prototype tool based on symbolic execution. The program-specification formalism we consider is Reachability Logic, a language-independent alternative to Hoare logics. Reachability Logic has a sound and relatively complete deduction system that offers a lot of freedom to the user regarding the manner and order of rule application, but it lacks a strategy for automatic proof construction. Hence, we propose a procedure for proof construction, in which symbolic execution plays a major role. We prove that, under reasonable conditions on its inputs (the operational semantics of a programming language, and a specification of a program, both given as sets of Reachability Logic formulas) our procedure is partially correct: if it terminates it correctly answers (positively or negatively) to the question of whether the given program specification holds when executing the program according to the given semantics. Termination, of course, cannot be guaranteed, since program-verification is an undecidable problem; but it does happen if the provided set of goals includes enough information in order to be circularly provable (using each other as hypotheses). We introduce a prototype program-verification tool implementing our procedure in the K language-definition framework, and illustrate it by verifying nontrivial programs written in languages defined in K. With respect to earlier versions of this work from 2013, program verification is now presented as a procedure (instead of a proof system), which leads to a direct implementation in the new version of our prototype tool. We also have a new theoretical result: *weak completeness*, which says that a negative answers returned by the verification procedure imply the fact that that the program does not meet its specification. Finally, since Andrei Arusoaie's arrival in the Dreampal team as a postdoc (Nov 2014) we have started working on certifying our verification procedure in the Coq proof assistant.

5.5.4. Language Definitions as Rewrite Theories

In [8] we study the relationships between language definition frameworks (e.g., the K framework) and rewrite theories (e.g., as those embodied in the Maude tool). K is a formal framework for defining the operational semantics of programming languages. It includes software tools for compiling K language definitions to Maude rewrite theories, for executing programs in the defined languages based on the Maude rewriting engine, and for analyzing programs by adapting various Maude analysis tools. A recent extension to the K tool suite is an automatic transformation of language definitions that enables the symbolic execution of programs, i.e., the execution of programs with symbolic inputs. In this paper we investigate more particularly the theoretical relationships between K language definitions and their translations to Maude, between symbolic extensions of K definitions and their Maude encodings, and how the relations between K definitions and their symbolic extensions are reflected on their respective representations in Maude. These results show, in particular, how analyses performed with Maude tools can be formally lifted up to the original language definitions. The results presented in this paper provide the theoretical underpinnings for the current version of the K-Maude tool.

5.6. Hardware chain for partial reconfiguration

The cost overhead due to the use of a softcore processor (MicroBlaze) to drive dynamic reconfiguration led us to explore alternative solutions. The one we have adopted is the use of a dedicated hardware IP (that can be invoked by HoMade) to control and manage dynamic and partial reconfiguration. This approach has led us to develop a complete hardware chain for partial bitstreams reads and writes. The proposed architecture is based on an external memory controller (DDR3) whose role is to manage bitstreams transfers from and to the DDR. Bitstreams loading are managed by a HoMade instruction implemented in a dedicated IP that drives the ICAP interface to transfer data into the reconfigurable area through the physical ICAP. One of the most important performance criteria of dynamic and partial reconfiguration is the reconfiguration time, that we always try to reduce while taking into account the compromise cost / area, speed and power consumption. Preliminary results give a transfer rate exceeding 500 MB/s. Such a result is clearly promising, especially since our hardware reconfiguration chain is constructed to be easily adaptable to SPMD (multi HoMade) needing parallel partial reconfiguration. This work has been the subject of a first communication in the GDR / SOCSIP conference in Paris: 11, 12, 13 June 2014.

5.7. Generic pixel distribution for parallel video processing application

In the frame of the PhD thesis of Karim Ali, we exploited this year the usage of parallel architectures for real-time image/video processing applications. Our main concern was the data distribution according to the parallelism level and respecting real-time processing constraint. As a first step, we proposed a generic pixel distribution model to be used with different image/video applications. Several parameters in the model can be configured according to the required size of the distributed macro-block with the possibility to control the sliding step in both horizontal and vertical directions. We have implemented our architecture on the Xilinx Zynq ZC706 FPGA evaluation board for two applications: the video downscaler (1:16) and the convolution filter. The experimental results showed the low hardware cost of the solution and how flexible is the model to be configured for different distribution scenarios. The architecture and experimental results were published in a paper entitled "A Generic Pixel Distribution Architecture for Parallel Video Processing" at Reconfigurable computing and FPGA international conference (ReConFig) in December 2014, Cancun, Mexico [7].

As a next step, we will reduce the operating clock frequency to decrease the power consumption while increasing the number of processing elements in the parallel architecture to maintain the same performance results. In this way, we will obtain a set of different design points differ in (area, power, other factors) and the system will have the ability to adapt its structure by moving between different design points according to the available resources to keep the same performance measurements. Furthermore, we will target intelligent transportation system, specially dynamic obstacle detection and tracking for autonomous vehicle navigation in collaboration with NAVYA (http://navya-technology.com).

5.8. Massively Parallel Dynamically Reconfigurable Multi-FPGA

In the frame of the PhD thesis of Venkatasubramanian Viswanathan, we conceived and validated a massively parallel and dynamically reconfigurable execution model for next generation high performance embedded systems. We have designed a multi-FPGA platform in order to conceive the massively parallel dynamically reconfigurable execution model. We have used several IP cores developed during the first two years of my PhD in order to test and validate the proposed model. We have proposed a new parallel dynamic reconfiguration mechanism for our architecture. We use our parallel reconfiguration model to reconfigure a subset or several IPs in parallel. We have proposed a partial reconfiguration model for next generation 3D FPGAs well-traced on the execution model (SPMD) in order to reconfigure in parallel a subset of the computing nodes. Finally, we have used the PicoComputing platform as an example to validate our proposed execution and reconfiguration models.

In order to demonstrate various features of such an architecture, we have implemented a scalable distributed secure H.264 encoding application with a FMC based high-speed sFPDP (serial Front Panel Data Port) data acquisition protocol to capture RAW video data. The system has been implemented on 3 different FPGAs, respecting the SPMD execution model managing several input video sources in parallel. We have measured various performance metrics of the proposed massively parallel dynamically reconfigurable system and demonstrated several benefits. This work is going to be published in the FPGA 2015 conference as a poster titled "A Parallel And Scalable Multi-FPGA based Architecture for High Performance Applications" [13].

Later an ICAP controller was setup for dynamic partial reconfiguration in order to swap IPs during runtime on a single FPGA. We have used this IP along with the parallel communication feature of the multi-FPGA architecture, in order to broadcast a partial bitstream to all FPGAs at the same time and to do a parallel DPR in several FPGAs, thus emulating the reconfiguration model for next generation 3D FPGAs. These results represent a conceptual proof for a massively parallel dynamically reconfigurable next generation embedded computers that will use 3D PFGAs and reconfigure several logic layers in parallel.

5.9. HoMade-based MPSoC

The goal of this work is to build an MPSoC based on HoMade. The aimed system is a completely dynamically reconfigurable system. This mean that both the processing elements (HoMade) and the interconnection network are dynamically reconfigurable. The basic block in this system developped here is the interconnection network. It is a MIN (Multistage Interconnection Network) that would utilize oversizing techniques in order to reconfigure the network depending on the traffic.

5.10. Communication-Computation Overlap in Massively Parallel System-on-Chip

The Synchronous Communication Asynchronous Computation (SCAC) model is an execution model dedicated to the Massively Parallel System-on-Chip. This model proposes a novel processing paradigm, teh communication-computation overlap [17]. This concept does not only consider the programming level but also the implementation level. Using a decoupled control structure, the synchronous communication control is performed independently of the asynchronous computation control. Separating these two control phases allows the programmer to define programming strategies that overlap communication by computation to decrease the execution time.

To achieve this communication-computation overlap in SCAC architecture while avoiding the centralized control, in addition to the master controller, we define a second hierarchical control level, namely the slave controllers. The concept of this dual ontrol structure departs from the centralized configuration and instead of a uni-processor master controlling a set of parallel Processing Elements (PEs), the master cooperates with a grid of parallel slave controllers which supervises the activities of cluster of PEs. Based on this decoupled control structure, the programmer can manage the master-slave program to overlap communication by computation phase. Therefore, the basic idea to implement this paradigm is to divide the principal program into small blocks of parallel instructions, called Slave Program (SP), and send these blocks to the activated PEs of the system. Then, according to a predefined mask, the slave controllers send the begin execution orders. In parallel to computation, the slave controllers manage the synchronous inter-node communication. Distinguish communication from computation needs the separation of these two phases in different blocks. This repartition should be provided at programming level. Then, the overlapped execution of these blocks will be done in parallel according to the program description.

The aim of these last works is to define a new paradigm of a communication-computation overlap in massively parallel System-on-Chip. This paradigm allows to decrease the execution time of parallel programs using specific strategies in the programming level and a partially decoupled control system in the hardware level. The difficulty of implementing this paradigm lies in the coordination between the programming level and the architecture designing level in order to hide the communication cost.

DYLISS Project-Team

6. New Results

6.1. Highlights of the Year

Four PhD theses were defended this year. They evidenced that ASP-technologies are now mature enough to perform data integration of large-scale bio-molecular datasets: classification of families of proteins [10], reconstruction of regulatory networks [13], reconstruction of metabolic network [11], and modelling of the discrete dynamics of a signalling or a regulatory network [12]. Importantly, symbolic classification technics have been adapted to exhibit relevant biological features: we used both formal concept analysis and semantic-based analysis for sequence and network analysis.

6.2. Data integration

Participants: Jacques Nicolas, Charles Bettembourg, Jérémie Bourdon, Jeanne Cambefort, Marie Chevallier, Guillaume Collet, Olivier Dameron, Damien Eveillard, Julie Laniau, Sylvain Prigent, Anne Siegel, Valentin Wucher.

Pan-genomic metabolic network of *Ectocarpus siliculosus*: We introduced the first metabolic network for the non-classical species *E. Siliculosus*, called EctoGEM. The reconstruction process includes draft reconstruction based on sequence and functional annotation analysis. It is followed by a combinatorial gap-filling process using the Meneco software based on answer set programming, a semantic analysis of the completion and a manual curation. This reconstruction enables a better understanding of organism biology and a reannotation of its genome. [*J. Cambefort, G. Collet, O. Dameron, D. Eveillard, S. Prigent, A. Siegel*] [22], [11]

New insights on bacteria associated with brown algae As an application of our tools for the reconstruction of metabolic networks, we have contributed to the analysis of the genome of a bacteria which lives in symbiosis with brown algae by investigating candidates for metabolic exchanges between the bacteria and the algae. [*G. Collet, J. Cambefort, A. Siegel*] [19] [Online publication]

Modeling parsimonious putative regulatory networks We integrated heterogeneous information from two types of network predictions to determine a causal explanation for the observed gene co-expression. We modeled this integration as a combinatorial optimization problem. We demonstrated that this problem belongs to the NP-hard complexity class. We proposed an heuristic approach to have an approximate solution in a practical execution time. Our evaluation showed that the E.coli regulatory network resulting from the application of this method has higher accuracy than the putative one built with traditional tools. [*A. Aravena, A. Siegel, D. Eveillard*] [26] [Online publication]

Modeling of a gene network between mRNAs and miRNAs to predict gene functions involved in phenotypic plasticity in the pea aphid and non coding RNA in pea aphid During its PhD, V. Wucher has built the first network model of gene regulation by microRNAs in pea aphid. The thesis has studied the discrimination between embryos development towards either sexual or asexual reproduction types in the pea aphid *Acyrthosiphon pisum*, at the genomic level. The study of the post-transcriptional regulation network implies both the identification of regulated elements during embryogenesis and the identification of the interaction modules between microRNAs and mRNAs using formal concept analysis. It helps the understanding of regulation differences between sexual and asexual embryogenesis. Moreover, it is the first step towards the modeling of the entire set of genes regulations in pea aphid during embryogenesis. [*V. Wucher, J. Nicolas, F. Legeai (Genscale team)*] [13], [39], [30]

Using a large-scale knowledge database on reactions and regulations to exhibit key regulators A new formalism of regulated reactions combining biochemical transformations and regulatory effects was proposed to unify the different mechanisms contained in knowledge libraries. Based on a related causality graph, an algorithm was developed to propose a reasonable set of upstream regulators from lists of target molecules. Scores were added to candidates according to their ability to explain the greatest number of targets or only few specific ones. The method was validated on a real example related to glycolysis. [*P. Blavy, A. Siegel*] [18] [Online publication]

Semantic particularity measure for functional characterization of gene sets using gene ontology We propose a new approach to compute gene set particularities based on the information conveyed by Gene Ontology terms. A GO term informativeness can be computed using either its information content based on the term frequency in a corpus, or a function of the term's distance to the root. We demonstrated that the combination of semantic similarity and semantic particularity measures was able to identify genes with particular functions from among similar genes. This differentiation was not recognized using only a semantic similarity measure. [*C. Bettembourg, O. Dameron*] [17] [Online publication]

Integrating GALAXY workflows in a metadata management environment New tools are needed to enable the quick design and the intensive parallel execution of bioinformatics processes. Therefore, we proposed a new dataflow-oriented workflow management system dedicated to intensive bioinformatics tasks. We worked on the interoperatibility of bioinformatics workflows using a model-driven approach. Our results enable new import / export capabilities beetwen multiple workflow management environnements and insights to create a unique shared workflow model. [*O. Dameron, F. Moreews (Genscale team), Y. Le Bras (GenOuest platform), C. Monjeaud (GenOuest platform), O. Collin (GenOuest platform)*][36]

6.3. Time-series and asymptotic dynamics

Participants: Anne Siegel, Jérémie Bourdon, Jeanne Cambefort, Damien Eveillard, Vincent Picard, Nathalie Théret, Santiago Videla.

Reasoning on the response of logical signaling networks with boolean models A series of papers and a PhD thesis focused on modeling the response of logical signaling networks by means of automated reasoning using ASP. In this context, a crucial issue is automatic learning of logical networks from partial observations of input/output behaviours, in order to achieve unbiased and robust discoveries. Experiments showed that many networks can be compatible with a given set of experimental observations. In a review chapter, we first discuss how ASP can be used to exhaustively enumerate all these logical networks. Next, in order to gain control over the system, we look for intervention strategies that force a set of target species into a desired steady state. Finally, we discuss the usage of ASP for solving the aforementioned problems and the novelty of our approach with respect to existing methods. [*S. Videla, A. Siegel, J. Nicolas*] [23], [38], [12] [Online publication]

Integrative modeling framework for signaling networks based on guarded transitions models We develop a new non-ambiguous formal interpretation of signaling pathways as discrete dynamic models. The resulting language, Computer-Aided Design for BIOlogical Models (CADBIOM), is based on a simplified version of guarded transitions in which we introduced temporal parameters for each transition to manage competition and cooperation between parts of the models . Tools for simulation and model checking analyses using the formal Cadbiom language have been developed (http://cadbiom.genouest.org). Using CADBIOM, we built the first discrete model of TGF-β signaling networks by automatically integrating the 137 human signaling maps from the Pathway Interaction Database into a single unified dynamic model. Temporal property-checking analyses of 15934 trajectories that regulate 145 TGF-$beta$ target genes reveal the association of specific pathways with distinct biological processes. [*G Andrieux, M Le Borgne, N. Théret*] [15] [Online publication]

Exploring metabolism flexibility in complex organisms through quantitative study of precursor sets for system outputs We extended a Flux-Balanced-Analysis approach to quantify the precursor composition of each system output and to discuss the biological relevance of a set of flux in a given metabolic network. The composition is called contribution of inputs over outputs [AIO]. In order to further investigate metabolic network flexibility, we have proposed an efficient local search algorithm computing the extremal values of

AIO coefficients. This approach enables to discriminate diets without making any assumption on the internal behaviour of the system. [*O. Abdou-Arbi, J. Bourdon, A. Siegel*] [14] [Online publication]

Multivariate Normal Approximation for the Stochastic Simulation Algorithm: limit theorem and applications We prove a central limit theorem on the asymptotic stochastic dynamical behavior of the outputs of a reaction network under certain steady-state assumptions. We present multiple applications including a constraints-based approach to verify asymptotic properties on the output moments without prior knowledge about the kinetic parameters. [*V. Picard, A. Siegel, J. Bourdon*] [33] [Online publication]

Numeric model for initiation of translation in sea-urchin We use a numeric-based modeling approach to study the regulation of protein synthesis following fertilization in sea urchin. This approach based on parcimonious modelling evidenced that two processes are required to explain experimental data: a destabilization of eIF4E:4E-BP complex and a great stimulation of the 4E-BP-degradation mechanism, both rapamycin-sensitive [*A. Siegel, J. Bourdon*] [21] [Online publication]

6.4. Sequence annotation

Participants: François Coste, Aymeric Antoine-Lorquin, Catherine Belleannée, Guillaume Collet, Gaëlle Garet, Clovis Galiez, Laurent Miclet, Jacques Nicolas, Valentin Wucher.

Automated Enzyme Classification by Formal Concept Analysis Guessing enzyme's functional activity from its sequence is a crucial task that can be approached by comparing the new sequences with those of already known enzymes labeled by a family class. This task is difficult because the activity is based on a combination of small sequence patterns and sequences greatly evolved over time. We have designed a classifier based on the identification of common subsequence blocks between known and new enzymes and the search of formal concepts built on the cross product of blocks and sequences for each class. Since new enzyme families may emerge, it is important to propose simultaneously a first classification of enzymes that cannot be assigned to a known family. Formal Concept Analysis offers a nice framework to set this task as an optimization problem on the set of concepts. The classifier has been tested with success on a particular set of enzymes present in a large variety of species, the haloacid dehalogenase (HAD) superfamily. [*F. Coste, G. Garet, J. Nicolas*] [28], [10]

A bottom-up efficient algorithm learning substitutable languages from positive examples Based on Harris's substitutability criterion, the recent definitions of classes of substitutable languages have led to interesting polynomial learnability results for expressive formal languages. These classes are also promising for practical applications: in natural language analysis, because definitions have strong linguisitic support, but also in biology for modeling protein families, as suggested in our previous study introducing the class of local substitutable languages. But turning recent theoretical advances into practice badly needs truly operable algorithms. We present here an efficient learning algorithm, motivated by intelligibility and parsing efficiency of the result, which directly reduces the positive sample into a small non redundant canonical grammar of the target substitutable language. Thanks to this new algorithm, we have been able to extend our experimentation to a complete protein dataset confirming that it is possible to learn grammars on proteins with high specificity and good sensitivity by a generalization based on local substitutability. [*F. Coste, G. Garet, J. Nicolas*] [29], [10]

Logol: Expressive Pattern Matching in sequences. Application to Ribosomal Frameshift Modeling Logol consists in both a language for describing biological patterns, and an associated parser for effective pattern search in sequences (RNA, DNA or protein). The Logol language, based on an high level grammatical formalism (String Variable Grammars), allows to express flexible patterns (with mispairings and indels) composed of both sequential elements (such as motifs) and structural elements (such as repeats or pseudoknots). Its expressive power allows the design of sophisticated patterns such as the signature of "-1 programmed ribosomal frameshifting" (PRF) events in messenger RNA sequences. A PRF signature is a complex model composed of a slippery site followed by a pseudoknot located in a specific part of the sequence, which provides a good illustration of the Logol language power. [*C. Belleannée, J. Nicolas, O. Sallou (GenOuest platform)*] [27] [Online publication]

Identifying distant homologous viral sequences in metagenomes using protein structure information It is estimated that marine viruses daily kill about 20% of the ocean biomass. Identifying them in water samples is thus a biological issue of great importance. The metagenomic approach for virus identification is a challenging task since their sequences carry a lot of mutations and are very difficuly to identify by standard homology searches. The PEPS VAG project aims at establishing a novel methodology that uses structures of proteins as extra-information in order to annotate metagenomes without relying on sequence homology. In the context of the first experiments made on the metagenome of station 23 of the TARA Ocean Project, we used the structures of capsid proteins to infer the sequence signature of their fold, in order to find them in the metagenome. This work presents the methodology, the first experiments and the on-going improvements. [*C. Galiez, F. Coste*] [35]

Computational Protein Design: trying an Answer Set Programming approach to solve the problem The problem of *Computational Protein Design* aims at finding the best protein conformation to perform a given task. This problem can be reduced to an optimization problem, looking for the minimum of an energy function depending on the amino-acid interactions in the protein. The CPD problem may be easily modeled as an ASP program but a practical implementation able to work on real-sized instances has never been published. We have raised the main source of difficulty for current ASP solvers and ran a series of benchmarks highlighting the importance of finding a good upper bound estimation of the target minimum energy to reduce the amount of combinatorial search. Our solution clearly outperforms a direct ASP implementation without this estimation and has comparable performances with respect to SAT-based approaches. It remains less efficient than a recent approach by cost function networks, showing there still exists some place for improving the optimization component in ASP with more dynamical strategies. [*J. Nicolas, H. Bazille*] [34]

Searching for Optimal Orders for Discretized Distance Geometry The Molecular Distance Geometry Problem (MDGP) is the problem of finding the possible conformations of a molecule by exploiting available information on some distances between pairs of its atoms. When some assumptions are satisfied, the MDGP can be discretized, so that the search domain of the problem becomes a tree where each node corresponds to a candidate position for an atom. The search tree can be efficiently explored by using an *interval* Branch & Prune (iBP) algorithm that can potentially enumerate all feasible conformations. In this context, the order given to the atoms of the molecule plays an important role, because it allows the discretization assumptions to be satisfied, and it also impacts the computational cost of the iBP algorithm. We have proposed a new discretization order for protein backbones based on the optimization of certain criteria for a faster exploration of the discretized search domain. To this aim, we express the search for optimal orders by a set of logical constraints in ASP. Our comparison with previously proposed orders for protein backbones shows that this new discretization order makes iBP perform better. [*J. Nicolas, A. Muccherino (Genscale Team)*] [43]

From analogical proportions in lattices to proportional analogies in formal concepts We provided an attempt at bridging formal concept analysis and the modeling of analogical proportions (i.e., statements of the form “a is to b as c is to d”). A suitable definition for analogical proportions in non distributive lattices is proposed and then applied to concept lattices. This enables us to compute what we call proportional analogies. In addition, we define the locally maximal subwords and locally minimal superwords common to a finite set of words. We also define the corresponding sets of alignments. We show that the constructed family of sets of alignments has the lattice structure. The study of analogical proportion in lattices gives hints to use this structure as a machine learning basis, aiming at inducing a generalization of the set of words. [*L. Miclet*] [32], [37]

DYOGENE Project-Team

6. New Results

6.1. Highlights of the Year

- F. Baccelli received 2014 IEEE Communications Society Stephen O. Rice Prize in the Field of Communications Theory:

  http://www.comsoc.org/about/memberprograms/comsoc-awards/rice.
- F. Baccelli received 2014 IEEE Communications Society Leonard G. Abraham Prize in the Field of Communications Systems:

  http://www.comsoc.org/about/memberprograms/comsoc-awards/abraham.
- F. Baccelli received ACM Sigmetrics Achievement Award 2014:

  http://www.sigmetrics.org/achievementaward-2014.shtml.
- F. Simatos received 2014 ACM SIGMETRICS Rising Star Researcher Award:

  http://www.sigmetrics.org/risingstar-2014.shtml.
- P. Brémaud published a book "Fourier Analysis and Stochastic Processes". Series: Universitext. Springer, Sept. 2014 - 385 pages.
- PhD student C. Rovetta received best tool paper award at Valuetools 2014 for the paper [18].

6.2. On Spatial Point Processes with Uniform Births and Deaths by Random Connection

With I. Norros (VTT Finland) and F. Mathieu (Bell Labs France), F. Baccelli has continued the line of thought on the geometry of Peer-to-Peer systems that was initiated in their Infocom 13 paper. This type of dynamics leads to a class of spatial birth and death process of the Euclidean space where the birth rate is constant and the death rate of a given point is the shot noise created at its location by the other points of the current configuration for some response function f. An equivalent view point is that each pair of points of the configuration establishes a random connection at an exponential time determined by f, which results in the death of one of the two points. The research concentrated on space-motion invariant processes of this type. Under some natural conditions on f, one can construct the unique time-stationary regime of this class of point processes by a coupling argument. The birth and death structure can then be used to establish a hierarchy of balance integral relations between the factorial moment measures. One can also show that the time-stationary point process exhibits a certain kind of repulsion between its points that is called f-repulsion.

These results were published in [29].

6.3. A Stochastic Geometry Framework for Analyzing Pairwise-Cooperative Cellular Networks

With A. Giovanidis, IMT, F. Baccelli has studied a cooperation model where the positions of base stations follow a Poisson point process distribution and where Voronoi cells define the planar areas associated with them. For the service of each user, either one or two base stations are involved. If two, these cooperate by exchange of user data and reduced channel information (channel phase, second neighbour interference) with conferencing over some backhaul link. The total user transmission power is split between them and a common message is encoded, which is coherently transmitted by the stations. The decision for a user to choose service with or without cooperation is directed by a family of geometric policies. The suggested policies further control the shape of coverage contours in favor of cell-edge areas. Analytic expressions based on stochastic geometry are derived for the coverage probability in the network. Their numerical evaluation shows benefits from cooperation, which are enhanced when Dirty Paper Coding is applied to eliminate the second neighbour interference.

These results were published in [7].

6.4. Analysis of a Proportionally Fair and Locally Adaptive spatial Aloha in Poisson Networks

With C. Singh (IIT), F. Baccelli and B. Blaszczyszyn worked on combining adaptive protocol design, utility maximization and stochastic geometry. The focus was on a spatial adaptation of Aloha within the framework of ad hoc networks. Quasi-static networks are considered, in which mobiles learn the local topology and incorporate this information to adapt their medium access probability (MAP) selection to their local environment. The cases where nodes cooperate in a distributed way to maximize the global throughput or to achieve either proportional fair or max-min fair medium access were considered. The proportionally fair sharing case leads to closed-form performance expressions in two extreme cases: (1) the case without topology information, where the analysis boils down to a parametric optimization problem leveraging stochastic geometry; (2) the case with full network topology information, which was recently solved using shot-noise techniques. It was shown that there exists a continuum of adaptive controls between these two extremes, based on local stopping sets, which can also be analyzed in closed form. These control schemes are implementable, in contrast to the full information case which is not. As local information increases, the performance levels of these schemes are shown to get arbitrarily close to those of the full information scheme. The analytical results are combined with discrete event simulation to provide a detailed evaluation of the performance of this class of medium access controls.

These results were published in [16].

6.5. Quality of Real-Time Streaming in Wireless Cellular Networks - Stochastic Modeling and Analysis

We present a new stochastic service model with capacity sharing and interruptions, appropriate for the evaluation of the quality of real-time streaming (e.g. mobile TV) in wireless cellular networks [2]. It takes into account multi-class Markovian process of call arrivals (to capture different radio channel conditions, requested streaming bit-rates and call-durations) and allows for a general resource allocation policy saying which users are temporarily denied the requested fixed streaming bit-rates (put in outage) due to resource constraints. We develop general expressions for the performance characteristics of this model, including the mean outage duration and the mean number of outage incidents for a typical user of a given class, involving only the steady-state of the traffic demand. We propose also a natural class of least-effort-served-first resource allocation policies, which cope with optimality and fairness issues known in wireless networks, and whose performance metric! s can be easily calculated using Fourier analysis of Poisson variables. We specify and use our model to analyze the quality of real time streaming in 3GPP Long Term Evolution (LTE) cellular networks. Our results can be used for the dimensioning of these networks.

6.6. On Comparison of Clustering Properties of Point Processes

In [3], we propose a new comparison tool for spatial homogeneity of point processes, based on the joint examination of void probabilities and factorial moment measures. We prove that determinantal and permanental processes, as well as, more generally, negatively and positively associated point processes are comparable in this sense to the Poisson point process of the same mean measure. We provide some motivating results on percolation and coverage processes, and preview further ones on other stochastic geometric models, such as minimal spanning forests, Lilypond growth models, and random simplicial complexes, showing that the new tool is relevant for a systemic approach to the study of macroscopic properties of non-Poisson point processes. This new comparison is also implied by the directionally convex ordering of point processes, which has already been shown to be relevant to the comparison of the spatial homogeneity of point processes. For this latter ordering, using a notion of lattice perturbation, we provide a large monotone spectrum of comparable point processes, ranging from periodic grids to Cox processes, and encompassing Poisson point processes as well. They are intended to serve as a platform for further theoretical and numerical studies of clustering, as well as simple models of random point patterns to be used in applications where neither complete regularity nor the total independence property are realistic assumptions.

6.7. SINR in Wireless Networks and the Two-Parameter Poisson-Dirichlet Process

Stochastic geometry models of wireless networks based on Poisson point processes are increasingly being developed with a focus on studying various signal-to-interference-plus-noise ratio (SINR) values. In [9], we show that the SINR values experienced by a typical user with respect to different base stations of a Poissonian cellular network are related to a specific instance of the so-called two-parameter Poisson-Dirichlet process. This process has many interesting properties as well as applications in various fields. We give examples of several results proved for this process that are of immediate or potential interest in the development of analytic tools for cellular networks. Some of them simplify or are akin to certain results that are being developed in the network literature. By doing this we hope to motivate further research and use of Poisson-Dirichlet processes in this new setting.

6.8. How User Throughput Depends on the Traffic Demand in Large Cellular Networks

In [17], we assume a space-time Poisson process of call arrivals on the infinite plane, independently marked by data volumes and served by a cellular network modeled by an infinite ergodic point process of base stations. Each point of this point process represents the location of a base station that applies a processor sharing policy to serve users arriving in its vicinity, modeled by the Voronoi cell, possibly perturbed by some random signal propagation effects. User service rates depend on their signal-to-interference-and-noise ratios with respect to the serving station. Little's law allows to express the mean user throughput in any region of this network model as the ratio of the mean traffic demand to the steady-state mean number of users in this region. Using ergodic arguments and the Palm theoretic formalism, we define a global mean user throughput in the cellular network and prove that it is equal to the ratio of mean traffic demand to the mean number of users in the steady st! ate of the "typical cell" of the network. Here, both means account for double averaging: over time and network geometry, and can be related to the per-surface traffic demand, base-station density and the spatial distribution of the signal-to-interference-and-noise ratio. This latter accounts for network irregularities, shadowing and cell dependence via some cell-load equations. Inspired by the analysis of the typical cell, we propose also a simpler, approximate, but fully analytic approach, called the mean cell approach. The key quantity explicitly calculated in this approach is the cell load. In analogy to the load factor of the (classical) M/G/1 processor sharing queue, it characterizes the stability condition, mean number of users and the mean user throughput. We validate our approach comparing analytical and simulation results for Poisson network model to real-network measurements.

6.9. Pioneers of Influence Propagation in Social Networks

In [20], we present a diffusion model developed by enriching the generalized random graph (a.k.a. configuration model), motivated by the phenomenon of viral marketing in social networks. The main results on this model are rigorously proved in [3], and in this paper we focus on applications. Specifically, we consider random networks having Poisson and Power Law degree distributions where the nodes are assumed to have varying attitudes towards influence propagation, which we encode in the model by their transmitter degrees. We link a condition involving total degree and transmitter degree distributions to the effectiveness of a marketing campaign. This suggests a novel approach to decision-making by a firm in the context of viral marketing which does not depend on the detailed information of the network structure.

6.10. QoS and Network Performance Estimation in Heterogeneous Cellular Networks Validated by Real-Field Measurements

Mobile network operators observe a significant disparity of quality of service (QoS) and network performance metrics, such as the mean user throughput, the mean number of users and the cell load, over different network base stations. The principal reason being the fact that real networks are never perfectly hexagonal, base stations are subject to different radio conditions, and may have different engineering parameters. In [21], we propose a model that takes into account these network irregularities in a probabilistic manner, in particular assuming Poisson spatial location of base stations, lognormal shadowing and random transmission powers. Performance of base stations is modeled by spatial processor sharing queues, which are made dependent of each other via a system of load equations. In order to validate our approach, we estimate all the model parameters from the data collected in a commercial network, solve it and compare the spatial variability of the QoS and performance metrics! in the model to the real network performance metrics. Considering two scenarios: downtown of a big city and a mid-size city, we show that our model predicts well the network performance.

6.11. Clustering Comparison of Point Processes with Applications to Random Geometric Models

In [27], we review some examples, methods, and recent results involving comparison of clustering properties of point processes. Our approach is founded on some basic observations allowing us to consider void probabilities and moment measures as two complementary tools for capturing clustering phenomena in point processes. As might be expected, smaller values of these characteristics indicate less clustering. Also, various global and local functionals of random geometric models driven by point processes admit more or less explicit bounds involving void probabilities and moment measures, thus aiding the study of impact of clustering of the underlying point process. When stronger tools are needed, directional convex ordering of point processes happens to be an appropriate choice, as well as the notion of (positive or negative) association, when comparison to the Poisson point process is considered. We explain the relations between these tools and provide examples of point processes admitting them. Furthermore, we sketch some recent results obtained using the aforementioned comparison tools, regarding percolation and coverage properties of the germ-grain model, the SINR model, subgraph counts in random geometric graphs, and more generally, U-statistics of point processes. We also mention some results on Betti numbers for Čech and Vietoris-Rips random complexes generated by stationary point processes. A general observation is that many of the results derived previously for the Poisson point process generalise to some “sub-Poisson” processes, defined as those clustering less than the Poisson process in the sense of void probabilities and moment measures, negative association or dcx-ordering.

6.12. Sublinear-Time Algorithms for Monomer-Dimer Systems on Bounded Degree Graphs

For a graph G, let $Z(G,\lambda)$ be the partition function of the monomer-dimer system defined by $\sum_k m_k(G)\lambda^k$, where $m_k(G)$ is the number of matchings of size k in G. In [11], we consider graphs of bounded degree and develop a sublinear-time algorithm for estimating $\log Z(G,\lambda)$ at an arbitrary value $\lambda > 0$ within additive error ϵn with high probability. The query complexity of our algorithm does not depend on the size of G and is polynomial in $1/\epsilon$, and we also provide a lower bound quadratic in $1/\epsilon$ for this problem. This is the first analysis of a sublinear-time approximation algorithm for a $\#P$-complete problem. Our approach is based on the correlation decay of the Gibbs distribution associated with $Z(G,\lambda)$. We show that our algorithm approximates the probability for a vertex to be covered by a matching, sampled according to this Gibbs distribution, in a near-optimal sublinear time. We extend our results to approximate the average size and the entropy of such a matching within an additive error with high probability, where again the query complexity is polynomial in $1/\epsilon$ and the lower bound is quadratic in $1/\epsilon$. Our algorithms are simple to implement and of practical use when dealing with massive datasets. Our results extend to other systems where the correlation decay is known to hold as for the independent set problem up to the critical activity.

6.13. How Clustering Affects Epidemic in Random Networks

Motivated by the analysis of social networks, we study a model of random networks that has both a given degree distribution and a tunable clustering coefficient. We consider two types of growth processes on these graphs: diffusion and symmetric threshold model. The diffusion process is inspired from epidemic models. It is characterized by an infection probability, each neighbor transmitting the epidemic independently. In the symmetric threshold process, the interactions are still local but the propagation rule is governed by a threshold (that might vary among the different nodes). An interesting example of symmetric threshold process is the contagion process, which is inspired by a simple coordination game played on the network. Both types of processes have been used to model spread of new ideas, technologies, viruses or worms and results have been obtained for random graphs with no clustering. In [6], we are able to analyze the impact of clustering on the growth processes. While clustering inhibits the diffusion process, its impact for the contagion process is more subtle and depends on the connectivity of the graph: in a low connectivity regime, clustering also inhibits the contagion, while in a high connectivity regime, clustering favors the appearance of global cascades but reduces their size. For both diffusion and symmetric threshold models, we characterize conditions under which global cascades are possible and compute their size explicitly, as a function of the degree distribution and the clustering coefficient. Our results are applied to regular or power-law graphs with exponential cutoff and shed new light on the impact of clustering.

6.14. Edge Label Inference in Generalized Stochastic Block Model: From Spectral Theory to Impossibility Results

The classical setting of community detection consists of networks exhibiting a clustered structure. To more accurately model real systems we consider a class of networks (i) whose edges may carry labels and (ii) which may lack a clustered structure. Specifically we assume that nodes possess latent attributes drawn from a general compact space and edges between two nodes are randomly generated and labeled according to some unknown distribution as a function of their latent attributes. Our goal is then to infer the edge label distributions from a partially observed network. In [22], we propose a computationally efficient spectral algorithm and show it allows for asymptotically correct inference when the average node degree could be as low as logarithmic in the total number of nodes. Conversely, if the average node degree is below a specific constant threshold, we show that no algorithm can achieve better inference than guessing without using the observations. As a byproduct of our analysis, we show that our model provides a general procedure to construct random graph models with a spectrum asymptotic to a pre-specified eigenvalue distribution such as a power-law distribution.

6.15. Balanced Graph Edge Partition

Balanced edge partition has emerged as a new approach to partition an input graph data for the purpose of scaling out parallel computations, which is of interest for several modern data analytics computation platforms, including platforms for iterative computations, machine learning problems, and graph databases. This new approach stands in a stark contrast to the traditional approach of balanced vertex partition, where for given number of partitions, the problem is to minimize the number of edges cut subject to balancing the vertex cardinality of partitions.

In [19], we first characterize the expected costs of vertex and edge partitions with and without aggregation of messages, for the commonly deployed policy of placing a vertex or an edge uniformly at random to one of the partitions. We then obtain the first approximation algorithms for the balanced edge-partition problem which for the case of no aggregation matches the best known approximation ratio for the balanced vertex-partition problem, and show that this remains to hold for the case with aggregation up to factor that is equal to the maximum in-degree of a vertex. We report results of an extensive empirical evaluation on a set of real-world graphs, which quantifies the benefits of edge- vs. vertex-partition, and demonstrates efficiency of natural greedy online assignments for the balanced edge-partition problem with and with no aggregation.

6.16. Streaming, Memory-limited Algorithms for Community Detection

In [23], we consider sparse networks consisting of a finite number of non-overlapping communities, i.e. disjoint clusters, so that there is higher density within clusters than across clusters. Both the intra- and inter-cluster edge densities vanish when the size of the graph grows large, making the cluster reconstruction problem nosier and hence difficult to solve. We are interested in scenarios where the network size is very large, so that the adjacency matrix of the graph is hard to manipulate and store. The data stream model in which columns of the adjacency matrix are revealed sequentially constitutes a natural framework in this setting. For this model, we develop two novel clustering algorithms that extract the clusters asymptotically accurately. The first algorithm is *offline*, as it needs to store and keep the assignments of nodes to clusters, and requires a memory that scales linearly with the network size. The second algorithm is *online*, as it may classify a node when the corresponding column is revealed and then discard this information. This algorithm requires a memory growing sub-linearly with the network size. To construct these efficient streaming memory-limited clustering algorithms, we first address the problem of clustering with partial information, where only a small proportion of the columns of the adjacency matrix is observed and develop, for this setting, a new spectral algorithm which is of independent interest.

6.17. State Space Collapse for Critical Multistage Epidemics

We study a multistage epidemic model which generalizes the SIR model and where infected individuals go through $K > 0$ stages of the epidemic before being removed. An infected individual in stage k may infect a susceptible individual, who directly goes to stage k of the epidemic; or it may go to the next stage $k + 1$ of the epidemic. For this model, we identify the critical regime in which we establish diffusion approximations. Surprisingly, the limiting diffusion exhibits an unusual form of state space collapse which we analyze in detail.

6.18. Perfect Sampling for Closed Queueing Networks

In [4], we investigate coupling from the past (CFTP) algorithms for closed queueing networks. The stationary distribution has a product form only in a very limited number of particular cases when queue capacity is finite, and numerical algorithms are intractable due to the cardinality of the state space. Moreover, closed networks do not exhibit any monotonic property enabling efficient CFTP. We derive a bounding chain for the CFTP algorithm for closed queueing networks. This bounding chain is based on a compact representation of sets of states that enables exact sampling from the stationary distribution without considering all initial conditions in the CFTP. The coupling time of the bounding chain is almost surely finite, and numerical experiments show that it is close to the coupling time of the exact chain.

In [18], we present Clones, a Matlab toolbox for exact sampling from the stationary distribution of a closed queueing net-work with finite capacities. This toolbox is based on recent results using a compact representation of sets of states that enables exact sampling from the stationary distribu-tion without considering all initial conditions in the coupling from the past (CFTP) scheme. This representation reduces the complexity of the one-step transition in the CFTP al-gorithm to O(KM 2), where K is the number of queues and M the total number of customers; while the cardinality of the state space is exponential in the number of queues. In this paper, we focus on the algorithmic and implementation issues. We propose a new representation, that leads to one-step transition complexity of the CFTP algorithm that is in O(KM). We provide a detailed description of our matrix-based implementation. The toolbox can be downloaded at http://www.di.ens.fr/~rovetta/Clones.

6.19. Individual Risk in Mean-Field Control Models for Decentralized Control, with Application to Automated Demand Response

Flexibility of energy consumption can be harnessed for the purposes of ancillary services in a large power grid. In prior work by the authors a randomized control architecture is introduced for individual loads for this purpose. In examples it is shown that the control architecture can be designed so that control of the loads is easy at the grid level: Tracking of a balancing authority reference signal is possible, while ensuring that the quality of service (QoS) for each load is acceptable on average. The analysis was based on a mean field limit (as the number of loads approaches infinity), combined with an LTI-system approximation of the aggregate nonlinear model. In [15], we examine in depth the issue of individual risk in these systems. The main contributions of the paper are of two kinds: Risk is modeled and quantified: (i) The average performance is not an adequate measure of success. It is found empirically that a histogram of QoS is approximately Gaussian, and consequently each load will eventually receive poor service. (ii) The variance can be estimated from a refinement of the LTI model that includes a white-noise disturbance; variance is a function of the randomized policy, as well as the power spectral density of the reference signal. Additional local control can eliminate risk: (iii) The histogram of QoS is truncated through this local control, so that strict bounds on service quality are guaranteed. (iv) This has insignificant impact on the grid-level performance, beyond a modest reduction in capacity of ancillary service.

6.20. Passive Dynamics in Mean Field Control

Mean-field models are a popular tool in a variety of fields. They provide an understanding of the impact of interactions among a large number of particles or people or other "self-interested agents", and are an increasingly popular tool in distributed control. In [14], we consider a particular randomized distributed control architecture introduced in our own recent work. In numerical results it was found that the associated mean-field model had attractive properties for purposes of control. In particular, when viewed as an input-output system, its linearization was found to be minimum phase. In this paper we take a closer look at the control model. The results are summarized as follows: (i) The Markov Decision Process framework of Todorov is extended to continuous time models, in which the "control cost" is based on relative entropy. This is the basis of the construction of a family of controlled Markovian generators. (ii) A decentralized control architecture is proposed in which each agent evolves as a controlled Markov process. A central authority broadcasts a common control signal to each agent. The central authority chooses this signal based on an aggregate scalar output of the Markovian agents. (iii) Provided the control-free system is a reversible Markov process, the following identity holds for the linearization,

$$\mathrm{Real}(G(j\omega)) = \mathrm{PSD}_Y(\omega) \geq 0 \qquad \omega \in \mathbb{R}\,,,$$

where the right hand side denotes the power spectral density for the output of any one of the individual (control-free) Markov processes.

6.21. Optimization of Dynamic Matching Models

The bipartite matching model was born in the work of Gale and Shapley, who proposed the stable marriage problem in the 1960s. In [36], we consider a dynamic setting, modeled as a multi-class queueing network or MDP model. The goal is to compute a policy for the matching model that is optimal in the average cost sense. Computation of an optimal policy is not possible in general, but we obtain insight by considering relaxations. The main technical result is a form of "heavy traffic" asymptotic optimality. For a parameterized family of models in which the network load approaches capacity, a variant of the MaxWeight policy is approximately optimal, with bounded regret, even though the average cost grows without bound. Numerical results demonstrate that the policies introduced in this paper typically have much lower cost when compared to polices considered in prior work.

6.22. Stochastic Bounds with a Low Rank Decomposition

In [5], we investigate how we can bound a discrete time Markov chain (DTMC) by a stochastic matrix with a low rank decomposition. We show how the complexity of the analysis for steady-state and transient distributions can be simplified when we take into account the decomposition. Finally, we show how we can obtain a monotone stochastic upper bound with a low rank decomposition.

6.23. Generalizations of Bounds on the Index of Convergence to Weighted Digraphs

Sequences of maximum-weight walks of a growing length in weighted digraphs have many applications in manufacturing and transportation systems, as they encode important performance parameters. It is well-known that they eventually enter a periodic regime if the digraph is strongly connected. The length of their transient phase depends, in general, both on the size of digraph and on the magnitude of the weights. In this paper, we show that certain bounds on the transients of unweighted digraphs, such as the bounds of Wielandt, Dulmage-Mendelsohn, Schwarz, Kim, and Gregory-Kirkland-Pullman, remain true for critical nodes in weighted digraphs.

This work was done by Thomas Nowak together with Glenn Merlet from Aix-Marseille Unversité, Hans Schneider from the University of Wisconsin at Madison, and Sergeĭ Sergeev from the University of Birmingham. It was presented at the 53th IEEE Conference on Decision and Control and appeared in the journal Discrete Applied Mathematics.

6.24. Approximate Consensus in Highly Dynamic Networks: The Role of Averaging Algorithms

In this paper, we investigate the approximate consensus problem in highly dynamic networks in which topology may change continually and unpredictably. We prove that in both synchronous and partially synchronous systems, approximate consensus is solvable if and only if the communication graph in each round has a rooted spanning tree, i.e., there is a coordinator at each time. The striking point in this result is that the coordinator is not required to be unique and can change arbitrarily from round to round. Interestingly, the class of averaging algorithms which are memoryless and require no process identities entirely captures the solvability issue of approximate consensus in that the problem is solvable if and only if it can be solved using any averaging algorithm. Concerning the time complexity of averaging algorithms, we show that approximate consensus can be achieved with precision of ε in a coordinated network model in $O(n^{n+1} \log 1/\varepsilon)$ synchronous rounds, and in $O((\Delta n)^{n\Delta+1} \log 1/\varepsilon)$ rounds when the maximum round delay for a message to be delivered is Δ. We investigate various network models in which this exponential bound in the number of nodes reduces to a polynomial bound, and we prove that a general upper bound on the time complexity of averaging algorithms has to be exponential. We apply our results to networked systems with a fixed topology and classical benign fault models, and deduce both known and new results for approximate consensus in these systems. In particular, we show that for solving approximate consensus, a complete network can tolerate up to $2n-3$ arbitrarily located link faults at every round, in contrast with the impossibility result established by Santoro and Widmayer (STACS '89) showing that exact consensus is not solvable with $n-1$ link faults per round originating from the same node.

This work was done by Thomas Nowak together with Bernadette Charron-Bost from the CNRS and Matthias Függer from Vienna University of Technology. It is currently under submission.

6.25. Towards Binary Circuit Models That Faithfully Capture Physical Solvability

In contrast to analog models, binary circuit models are high-level abstractions that play an important role in assess-ing the correctness and performance characteristics of digital circuit designs: (i) modern circuit design relies on fast digital timing simulation tools and, hence, on binary-valued circuit models that faithfully model signal propagation, even throughout a complex design, and (ii) binary circuit models provide a level of abstraction that is amenable to formal correctness proofs. A mandatory feature of any such model is the ability to trace glitches and other short pulses precisely as they occur in physical circuits, as their presence may affect a circuit's correctness and its performance characteristics. Unfortunately, it was recently proved [Függer et al., ASYNC'13] that none of the existing binary-valued circuit models proposed so far, including the two most commonly used pure and inertial delay channels and any other bounded single-history channel, is realistic in the following sense: For the simple Short-Pulse Filtration (SPF) problem, which is related to a circuit's ability to suppress a single glitch, they showed that every bounded single-history channel either contradicts the unsolvability of SPF in bounded time or the solvability of SPF in unbounded time in physical circuits, i.e., no existing model correctly captures physical solvability with respect to glitch propagation. We propose a binary circuit model, based on so-called in-volution channels, which do not suffer from this deficiency. In sharp contrast to what is possible with all the existing models, they allow to solve the SPF problem precisely when this is possible in physical circuits. To the best of our knowledge, our involution channel model is hence the very first binary circuit model that realistically models glitch propagation, which makes it a promising candidate for developing more accurate tools for simulation and formal verification of digital circuits.

This work was done by Thomas Nowak together with Matthias Függer, Robert Najvirt, and Ulrich Schmid from Vienna University of Technolgy. It will be presented at the conference DATE 2105.

6.26. Weak CSR Expansions and Transience Bounds in Max-Plus Algebra

This paper aims to unify and extend existing techniques for deriving upper bounds on the transient of max-plus matrix powers. To this aim, we introduce the concept of weak CSR expansions: $A^t = CS^tR \oplus B^t$. We observe that most of the known bounds (implicitly) take the maximum of (i) a bound for the weak CSR expansion to hold, which does not depend on the values of the entries of the matrix but only on its pattern, and (ii) a bound for the CStR term to dominate. To improve and analyze (i), we consider various cycle replacement techniques and show that some of the known bounds for indices and exponents of digraphs apply here. We also show how to make use of various parameters of digraphs. To improve and analyze (ii), we introduce three different kinds of weak CSR expansions. As a result, we obtain a collection of bounds, in general incomparable to one another, but better than the bounds found in the literature.

This work was done by Thomas Nowak together with Glenn Merlet from Aix-Marseille Unversité and Sergeĭ Sergeev from the University of Birmingham. It appeared in the journal Linear Algebra and its Applications.

6.27. An Overview of Transience Bounds in Max-Plus Algebra

This book chapter surveys and discusses upper bounds on the length of the transient phase of max-plus linear systems and sequences of max-plus matrix powers. In particular, It explains how to extend a result by Nachtigall to yield a new approach for proving such bounds and states an asymptotic tightness result by using an example given by Hartmann and Arguelles.

This work was done by Thomas Nowak together with Bernadette Charron-Bost from the CNRS. It appeared in the book "Tropical and Idempotent Mathematics and Applications" in the AMS's book series Contemporary Mathematics.

E-MOTION Project-Team

5. New Results

5.1. Highlights of the Year

- C. Laugier, E. Mazer and K. Mekhnacha have been finalists for the Eurobotics Technology Award 2014. Title "Bayesian perception & Decision: from theory to industrial applications". March 2014.
- A. Nègre, L. Rummelhard, M. Perrollaz and C. Laugier had applied for a petenent "Procédé d'analyise d'une scene dybnamique, module d'analyse et programme d'ordinateur associés".

5.2. A new formulation of the Bayesian Occupancy Filter: a hybrid sampling based framework

Participants: Lukas Rummelhard, Amaury Nègre, Christian Laugier.

The Bayesian Occupancy Filter (BOF) is a discretized grid structure based bayesian algorithm, in which the environment is subdivised in cells to which random variables are linked. These random variables represent the state of occupancy and the motion field of the scene, without any notion of object detection and tracking, making the updating part of the filter an evaluation of the distribution of these variables, according to the new data acquisition. In the classic representation of the BOF, the motion field of each cell is represented as a neighborhood grid, the probability of the cell moving from the current one to another of the neighborhood being stocked in an histogram. If this representation is convenient for the update, since the potential antecedents of any cell is exactly determined by the structure, and so the propagation model is easily parallelizable, it also raises determinant issues :

- the structure requires the process rate to be constant, and a priori known.
- in the case of a moving grid, such as an application of car perception, many aliasing problems can appear, not only in the occupation grid, but in the motion fields of cells. A linear interpolation in 4-dimension field to fill each value of the histograms can quickly become unreasonable.
- to be able to match the slowest moves in the scene and the tiniest objects, the resolution of the grid and the motion histogram must be the high. On the other hand, since the system must be able to evaluate the speed of highly dynamic objects (typically, a moving car), the maximum encoded speed is to be high as well. This results in a necessary huge resolution grid, which prevent the system from being used with satisfying results on an embedded device. This huge grid is also mostly empty (most of the motion field histogram for a occupied cell will be empty). On top of that, the perception system being used to represent the direct environment of a moving car, the encoded velocity is a relative velocity, which implies, if we consider the maximal speed of a car to be V_max, to maintain a motion field able to represent speeds from $-2 * V_{max}$ to $2 * V_{max}$. The necessity of such a sized structure is a huge limitation of practical use of the method.

Considering those limitations, a new way to represent the motion field has been developped. To do so, a new formulation of the BOF has been elaborated. This new version allow to introduce in the filter itself a distinction between static and dynamic parts, and so adapt the computation power.

The Hybrid Sampling Bayesian Occupancy Filter (HSBOF) [21] is an evolution of the BOF, in which are introduced additionnal concepts and variables, such as probabilistic classification of the environment between static and dynamic areas, and adaptative motion model structure. The main idea of this new representation is to mix two forms of sampling of the surrounding :

- a uniform sampling, represented as a dense regular grid, for the static objects and the empty areas. In this part, only the occupancy is stored, as the motion model of the static part of the scene is inherent. In practice, the section of the environment includes the vast majority of the scene.

- a non uniform sampling, based on particles drawn in dynamic regions, allowing to focus the computational power on the estimation of their motion. The number of particles used to represent the motion of a particular cell is calculated according to various criterions, such as the confidence in the dynamism of the cell, in its estimated motion, the global needs in the scene, etc. Dynamic regions are resampled at every time step, the amount of particles associated to the different parts of the scene is dynamicly calculated.

The motion field in a cell is then represented as a set of samples from the distribution for values which are not null, and a weight given to the static hypothesis. The use of a set of samples to represent the motion field leads to a important decrease of the needed memory space, so do the classification between dynamic objects and static objects or free areas. In the updating process, the antecedent of a cell can be either from the static configuration or from the dynamic configuration, which are both way easier to project in the new reference frame of the moving grid: the static part requires a 2-dimension interpolation to be expressed in the new reference frame, the dynamic part a immediate particle association and a simple rotation of the velocity vectors.

This new version HSBOF is now used in the core of our systems in place of the previous version of the BOF. It presents important improvements in the quality of the estimations, while drastically reducing the memory and computation costs (easily by a 100 factor in term of memory).

5.2.1. *Probabilistic grid-based collision risk prediction*

Participants: Lukas Rummelhard, Amaury Nègre, Mathias Perrollaz, Christian Laugier.

We developped a new grid-based approach for collision risk prediction [23], based on the Hybrid-Sampling Bayesian Occupancy Filter framework. The idea is to compute an estimation of the Time To Contact (TTC) for each cell of the grid, instead of reasoning on objects. This strategy avoids to solve the difficult problem of multi-objects detection and tracking and provides a probabilistic estimation of the risk associated to each TTC value.

Using motion sensors embedded in the mobile robot (Inertial Measurement Unit, GPS, Wheel speed and steering sensor, visual odometry, etc.), the displacement of the grid between two updates is estimated. The full description of occupancy and dynamics of the scene given by the HSBOF is then used to assess collision risks in the future and even localize them in the grid. The risk evaluation consists in a short-term prediction of the scene configuration (figure 4 and of the robot position. This way a collision likelihood can be computed over time. Using those likelihoods, computed by cell and particle, an estimation of the risk over a period, and a localization of this risk in the grid are performed.

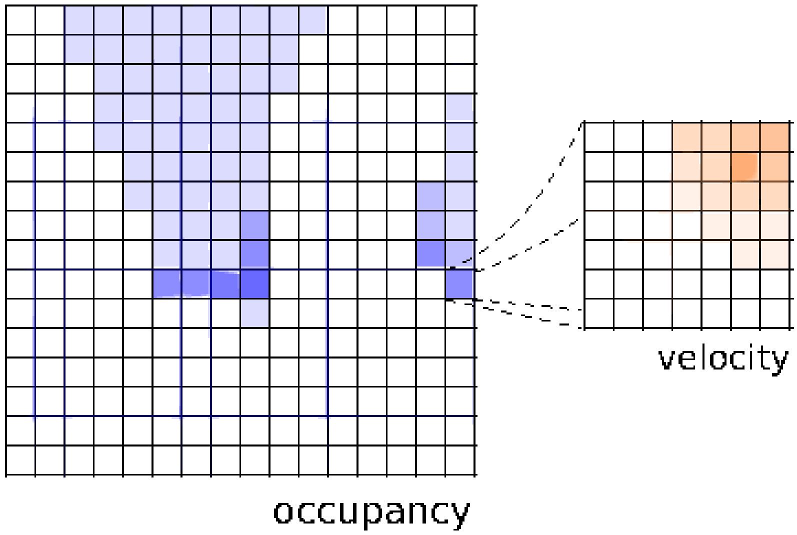
5.3. A new experimental platform for the Technological Research Institute (NanoElec)

Participants: Mathias Perrollaz, Nicolas Turro, Jean-François Cuniberto.

Within the framework of the PERFECT projet (founded by the IRT NanoElec), e-Motion has developped a new experimental platform, based on a Renault Zoe electrical vehicule (Fig. 7). This developpement takes advantage of the experience developped for creating the previous experimental plateform (a Lexus LS600H), and go further by integrating more sensors and more functionalities.

The vehicle is eqquipped with:

- 4 IBEO LUX laser scanners. Each of them scans 4 layers with a field of view of 85 degrees.
- one Velodyne HD64L 3D laser scnanner, capable of scanning 64 layers over 360 degrees.
- one trinocular stereo camera, Point Grey Bumblebee XB3, placed behind the windshield.
- 2 Ueye RGB cameras, looking forward and backward the vehicle.
- one XSens IMU/GPS sensor, used for positionning and ego-motion estimation.
- one ITRI 802.11p on-board unit, allowing V2X communication.



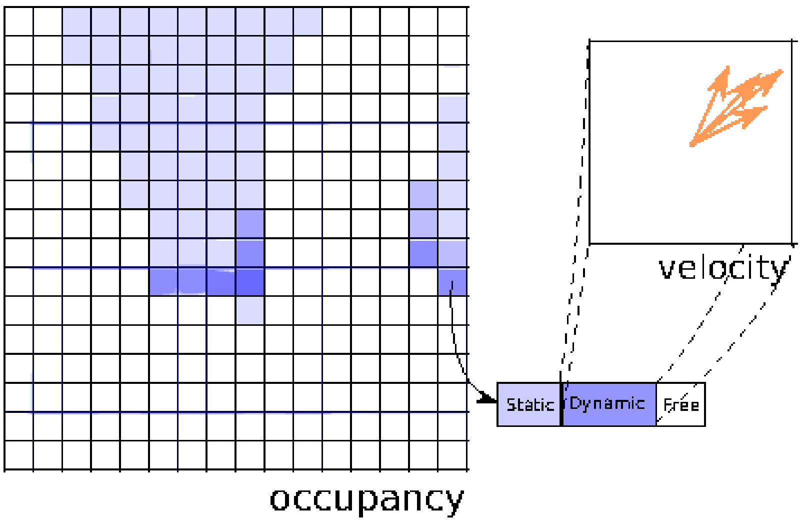


Figure 1.

Data representations in BOF and HSBOF formulation :
(a) Classic BOF representation : a 2 dimension grid, to each cell are assigned an occupancy value and a velocity histogram,
(b) Proposed representation : a 2 dimension grid, to each cell are assigned an occupancy value, a static coefficient $P(V=0)$ *and a set of particles drawn along* $P(V=v|V\neq 0)$

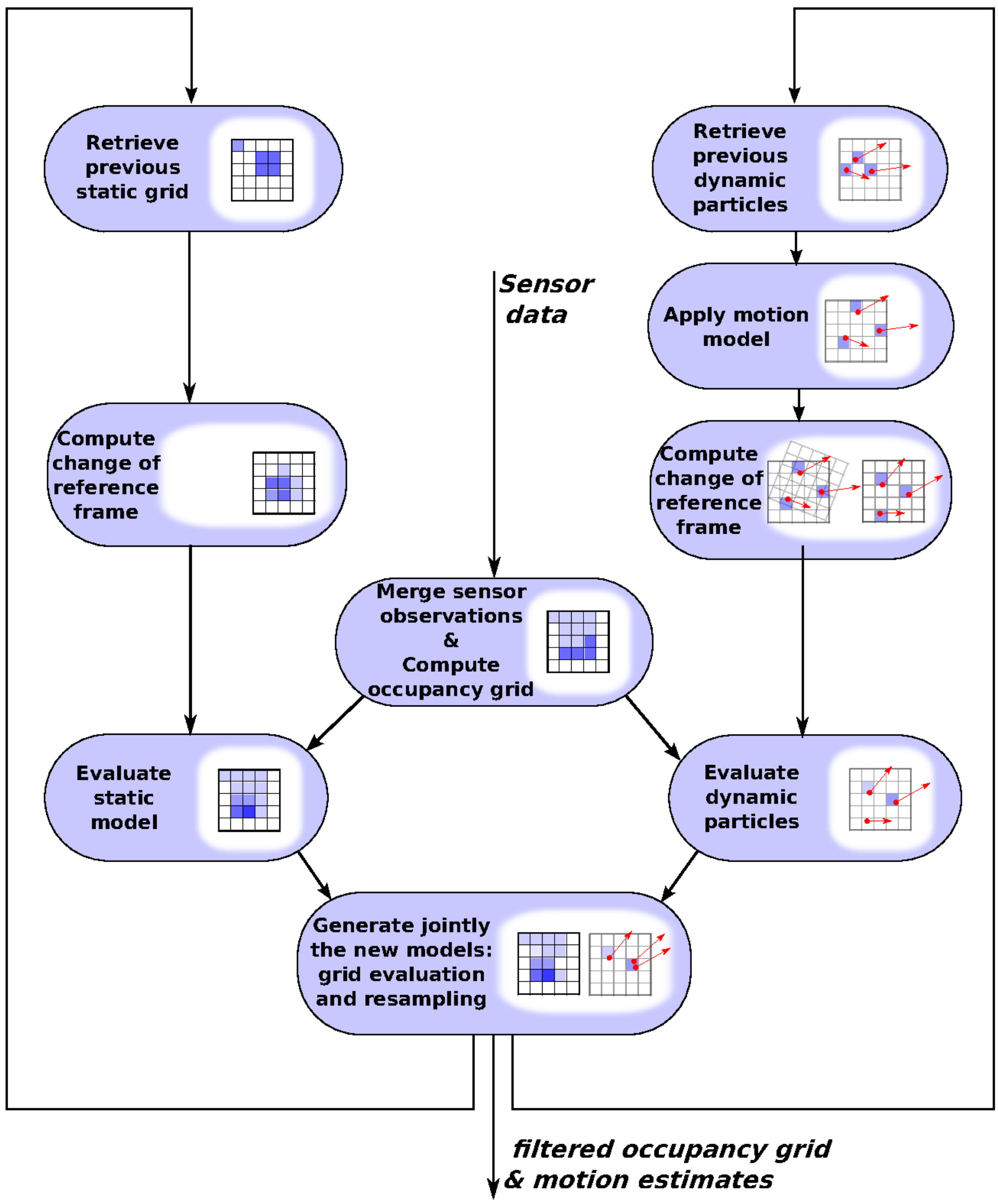


Figure 2. HSBOF algorithm summary. From sensor data instantaneous occupancy grids are successively computed. Those observations are integrated in a Bayesian filter in which coexist and jointly adapt two models, a static grid and a dynamic set of moving particles. The result is obtained by their combination, which provides a filtered occupancy grid as well as inferred motion distributions for cells.

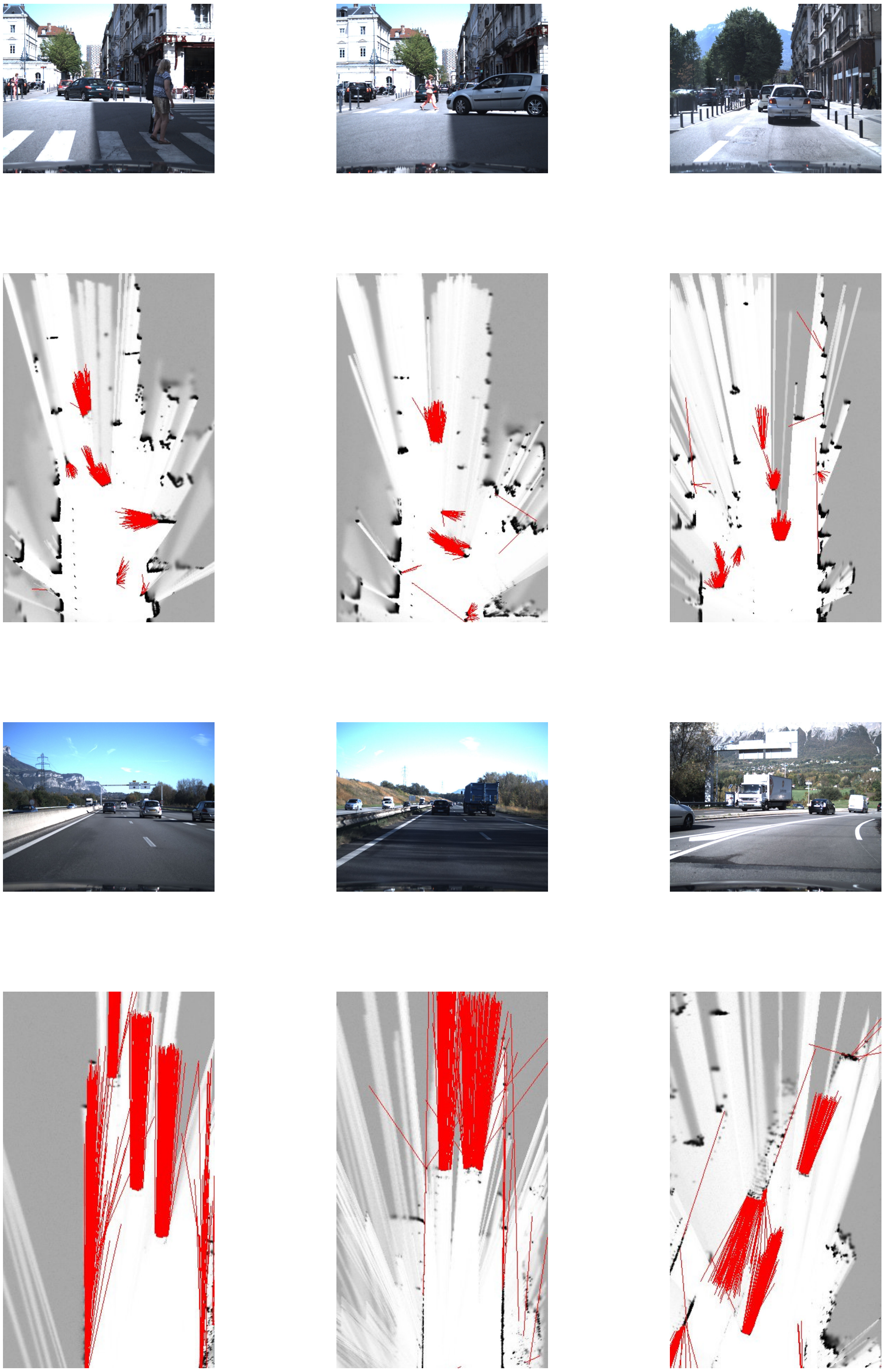


Figure 3. Resulting occupancy grid and velocity field on different urban and highway situations. White cells represent the free space, grey one the unknown space (hidden). Black cells represent the occupied space and red lines represent the average velocity vector for cell with a high dynamic probability.

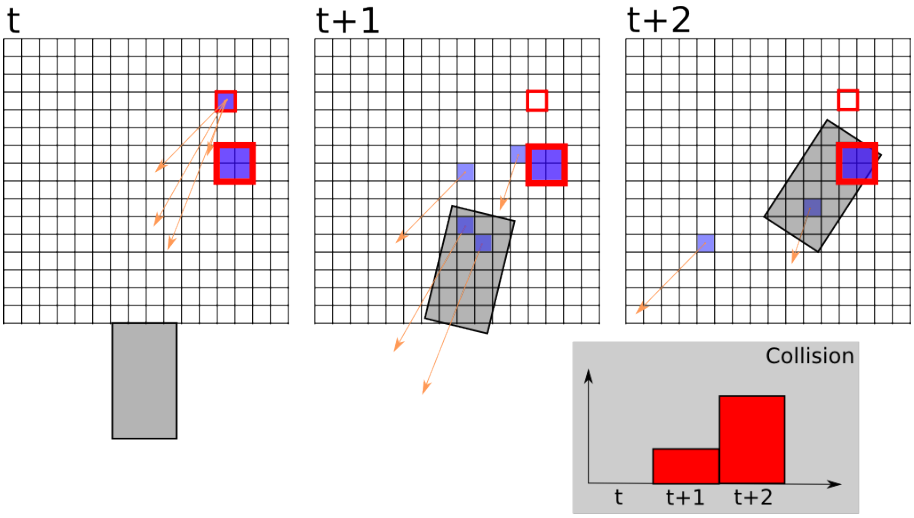


Figure 4. Collision risk estimation over time for a specific cell. The cell position is predicted according to its velocity, along with the mobile robot. This risk profile is computed for every cell, and then used to integrate over time the global collision risk.

(a) (b)

Figure 5. (a) Fake pedestrian used for experiments. (b) The mannequin is attached to a system with a runner, in order to allow lateral displacements.

(a)

(b)

(c)

(d)

Figure 6. Results of the system. Each image is a visual capture from the embedded camera, the estimated occupancy grid in front of the car (white for occupied, grey for unknown, black for empty), the estimated motion field (if a case is seen as dynamic, a red motion vector showing the average velocity in the cell is drawn on the map) and finally the estimated risk map for 0.5s. The first sequence (a) (b) presents the appearance of an occluded pedestrian, the second (c) (d) a moving pedestrian heading towards the road.

All the synchronization, display, play/record, and developpments capabilities are relying on the ROS middleware. The vehicle is fully operationnal at the end of 2014.
The vehicle is designed for experimenting in both ADAS (Advanced Driver Assistance Systems) and autonomous driving applications. In parallel, V2X communications are installed on the IRT "smart city" environment, so that the vehicle can evolve on this site and interact with it.

Figure 7. The Zoe experimental platform.

5.4. Visual localization with Open Street Map

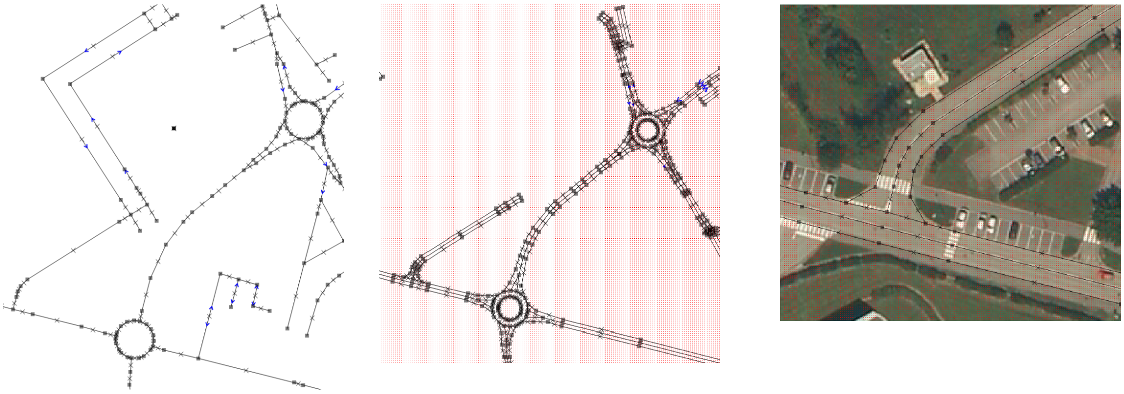
Participants: Jean-Alix David, Amaury Nègre.

Given the lack of precision of GPS for localization, it is necessary to implement new ways to improve localization. Here we introduce a new method using a geographic map and a camera to do so. The main point of this method is to combine sensor readings and known data about the environment. We detect lines on the road with the camera, and then compare the extracted lines to the ones stored in the map using ICP (Iterative Closest Point) algorithm.

The used map is OpenStreetMap, it allow to have information on the roads and lanes for example, but there is no information about white marking. So we generated semi-autonomously the lines given roads and number of lanes. Moreover we manually corrected the lines for crossroads using satellite image (see Figure 8).

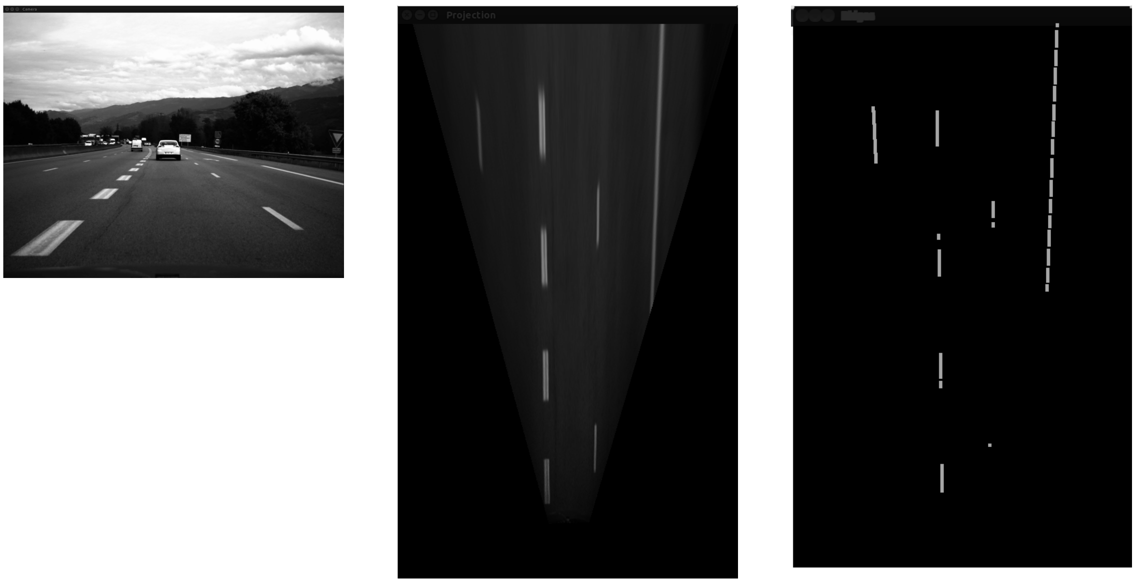
The line extraction is done using ridge detection on a top-down view of the camera image. Moreover we use GPU acceleration to improve performances during image processing (see Figure 9).

The OSM generated data and the lines extracted from the camera will then be matched and the transformation between the camera and the absolute map will be compute by using an Iterative Closest Point algorithm. In order to improve the precision, a bayesian filtering approach will also be used to merge the previous results with GPS and Inertial Measurement Unit data.



(a) Open Street Map data (b) Generated lines (c) Reprojected lines on aerial image after rectification

Figure 8. Semi-automatic lroad line generation from Open Street Map.



(a) (b) (c)

Figure 9. Line detection in camera images: projection in the ground plane (b) ridge extraction (c).

5.5. Human Centered Navigation in the physical world

5.5.1. Social Mapping

Participants: Panagiotis Papadakis, Anne Spalanzani, Christian Laugier.

With robots technology shifting towards entering human populated environments, the need for augmented perceptual robotic skills emerges that complement to human presence. In this integration, perception and adaptation to the implicit human social conventions plays a fundamental role. Toward this goal, we introduce in 2013 a novel methodology to detect and analyse complex spatial interactions of multiple people and encode them in the form of a social map, whose structure is obtained by computing a latent space representation of human proxemic behaviour [32]. In 2014, Panagiotis left to Lagadic-Sophia and we carried on this work by integrating a planning algorithm to validate the perception part on a real robot. This work was published at IROS 2014 [22].

5.5.2. Goal oriented risk based navigation in social and dynamic environment

Participants: Anne Spalanzani, Procopio Silveira-Stein, Gregoire Vignon, Christian Laugier.

Since 2008 we have proposed a new concept to integrate a probabilistic collision risk function linking planning and navigation methods with the perception and the prediction of the dynamic environments [31]. The likelihood of the obstacles' future trajectory and the probability of occupation are used to compute the risk of collision. A social filter was added to give the robot the ability to move in a social way (see Figure 10). In 2014, we obtained an Inria ADT(ADT PN2) to optimize and share the RiskRRT algorithm. This work is under development. We published in [15] a survey on human-aware navigation.

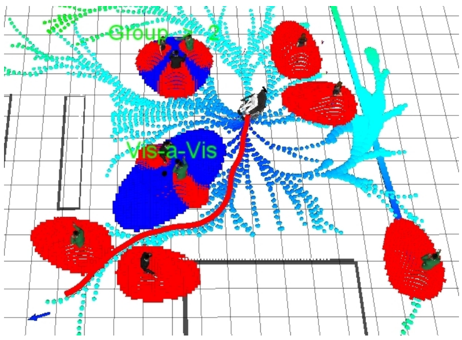


Figure 10. Illustration of the RiskRRT in a social environnment

5.5.3. Navigation Taking Advantage of Moving Agents

Participants: Procopio Silveira-Stein, Anne Spalanzani, Christian Laugier.

In this work, we proposes a different form of robotic navigation in dynamic environments, where the robot takes advantage of the motion of pedestrians, in order to improve its own navigation capabilities. The main idea is that, instead of treating persons as dynamic obstacles that should be avoided, they should be treated as special agents with an expert knowledge of navigating in dynamic scenarios. To benefit from the motion of pedestrians, this work proposes that the robot selects and follows them, so it can move along optimal paths, deviate from undetected obstacles, improve navigation in densely populated areas and increase its acceptance by other humans. In 2014, we focused on real experiments (see Figure 11 using the wheelchair and results were published in [16], [25], [26].

Figure 11. Switching navigation method between leader following and independent navi- gation. In image 1 the robot is engaged in leader following, while in the remaining it uses RiskRRT for the navigation.

5.5.4. Autonomous Wheelchair for Elders Assistance

Participants: Arturo Escobedo-Cabello, Gregoire Vignon, Anne Spalanzani, Christian Laugier.

The aging of world's population is bringing the need to provide robotic platforms capable to assist elder people to move [33]. It is necessary that such transportation is reliable, safe and comfortable. People with motor disabilities and elders are expected to benefit from new developments in the field of autonomous navigation robotics.

Autonomously driven wheelchairs are a real need for those patients who lack the strength or skills to drive a normal electric wheelchair. The services provided by this kind of robots can also be used to provide a service of comfort, assisting the user to perform difficult tasks as traversing a door, driving in a narrow corridor etc. In 2014, we combined user intention estimation, a navigation using social convention to perform comfortable trajectories (see Figure 12 . Results were published in the IROS conference [19]. Arturo Escobedo defended his phD in october 2014.

Figure 12. The Robotic wheelchair assists its user to acheive his task of navigation/.

5.6. Human modeling for situation understanding

5.6.1. Situation understanding and risk assessment for intelligent vehicles

Participants: Dizan Vasquez, Stéphanie Lefèvre, Suryansh Kumar, Yufeng Yu.

The work on this period has been aimed at establishing a solid theoretical and technological base for our research on situation understanding. A step in this direction was the elaboration of an in-depth survey of the current state of the art on the field, prepared together with the university of Berkeley [12]. In the framework of the same collaboration, we have been working on the introduction of human models in current Advanced Driving Assistance Systems (ADAS). This has led to the development of a novel Lane Keeping Assistance System (LKAS) which is able to learn the driver's driving patterns and use them to predict lane departures as well as to generate controls that mimic the driver's style and are, thus, deemed to be more acceptable. The approach has been evaluated against commercial LKAS using real field data, and the results show that the proposed approach is both more efficient and less intrusive than current approaches. This is, for the best of our knowledge the first use of human models withing LKAS and these results illustrate the strong potential that these models may have in ADAS.

Concerning autonomous navigation, we have focused on human-like motion planning for motion prediction. The main hypothesis is that people behave like planners whose motion optimizes some an unknown cost function. Under this assumption, the main challenge becomes to model that cost function and to learn its parameters from demonstrated behavior. This is called, depending on the community, either *Inverse Reinforcement Learning* (IRL) or *Inverse Optimal Control* (IOC). Now, a problem with IRL is that it requires examples of both desirable and undesirable behavior, which are difficult to obtain with a real platform. Additionally, there is no consistent benchmarking methodology to evaluate different approaches. This has motivated our work in a benchmark comprised of: (a) an evaluation methodology; (b) a simulated experimental platform (Fig. 13) based on the Torcs simulator ; and (c) real data gathered with our instrumented Lexus vehicle. The first prototype of this benchmark, developed together with students from Beijing University and IIIT Hyderabad, has been presented this year in a vehicular technologies conference .

Figure 13. Experimental platforms: left) Our Torcs-based racing simulator; right) sensor-equipped Lexus vehicle.

5.6.2. Socially compliant robot navigation in human environments

Participant: Dizan Vasquez.

The models we have applied to intelligent vehicles are also adapted in general to situations where mobile robots share their environment with humans. This has lead us to apply this techniques to the assistive robotics fields, given that it is one of e-motion's major applications axes. Our first effort in this sense has been to design and develop a robust experimental platform with baseline modules for motion planning, perception and social awareness .

In parallel we started working, in collaboration with the University of Freiburg, on a benchmarking platform for social compliant motion planning, close in spirit to the one proposed for intelligent vehicles. The platform (Fig. 14) is described in , it includes several motion planning and feature extraction algorithms as well as a pedestrian simulator based on Helbing's social force model .

5.7. Sensor Fusion for state parameters identification

Participants: Agostino Martinelli, Chiara Troiani.

5.7.1. General theoretical results

During this year we have focused our research on two distinct domains:

- the visual-inertial structure from motion problem;
- the derivation of analytical solutions for the probability distribution of a Brownian motion that satisfies the unicycle constraint.

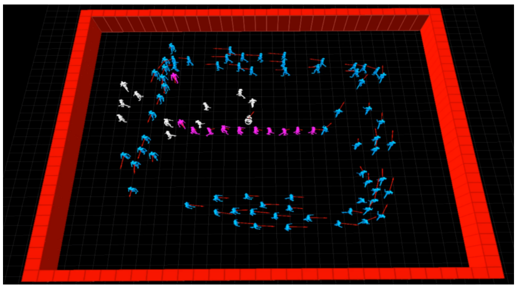


Figure 14. A screenshot of our pedestrian simulator.

The research carried out on the first domain is the follow up of our previous activity. We continued to investigate the observability properties of the visual inertial structure from motion and in particular we have analyzed the case when some of the inertial sensors are missing. This analysis has never been provided before and we started this investigation at the end of last year. During this year we confirmed the validity of our preliminary analysis and we also extended them. The preliminary results were obtained by referring to the case when at least five pont features are available and showed that the observability properties of visual inertial structure from motion do not change by removing all the three gyroscopes and one accelerometer. By removing a further accelerometer, if the camera is not extrinsically calibrated, the system loses part of its observability properties. On the other hand, if the camera is extrinsically calibrated, the system maintains the same observability properties as in the standard case. These results have been published on the journal Foundations and Trends in Robotics and have also been presented at the last ICRA conference[20].

We recently extended these results by considering the extreme case of a single point feature (i.e., not five). This analysis required to approach an open problem in control theory, called the Unknown Input Observability (UIO). In [20] we proposed a possible method to solve this UIO problem. However, we had to improve this method to deal with this extreme case (i.e., the case of one single point feature). Preliminary results on the extension of this method have been published as a research report [30] and we also plan to present them at the next American Control Conference. By applying this method to our problem, we obtained new interesting results. The new investigation allowed us to conclude that, even in the case of a single point feature, the information provided by a sensor suit composed by a monocular camera and two inertial sensors (along two independent axes and where at least one is an accelerometer) is the same as in the case of a complete inertial measurement unit (i.e., when the inertial sensors consist of three orthogonal accelerometers and three orthogonal gyroscopes). Our first objective is to validate these new results.

Regarding the second domain mentioned above, we have derived a complete analytical solution for the probability distribution of the configuration of a non-holonomic mobile robot that moves in two spatial dimensions by satisfying the unicycle kinematic constraints. The proposed solution differs from previous solutions since it is obtained by deriving the analytical expression of any-order moment of the probability distribution. To the best of our knowledge, an analytical expression for any-order moment that holds even in the case of arbitrary linear and angular speed, has never been derived before. To compute these moments, a direct integration of the Langevin equation has been carried out and each moment was expressed as a multiple integral of the deterministic motion (i.e., the known motion that would result in absence of noise).

For the special case when the ratio between the linear and angular speed is constant, the multiple integrals can be easily solved and expressed as the real or the imaginary part of suitable analytic functions. As an application of the derived analytical results, we also investigated the diffusivity of the considered Brownian motion for constant and for arbitrary time-dependent linear and angular speed. These results have been published on the journal of statistical mechanics [13] and also as a research report [29] where we added more specific considerations about the impact of the derived results on mobile robotics.

5.7.2. Applications with a Micro Aerial Vehicle

We continued our previous activity about the estimation of the relative motion between two consecutive camera views in order to introduce very efficient algorithms to remove the outliers of the feature-matching process. Thanks to their inherent efficiency, the proposed algorithms are very suitable for computationally-limited robots.

In particular, during this year, we extended the previous results by removing the assumption of planar motions. In this case, to obtain useful results, we had to include one more point feature (i.e., the proposed algorithms only use two feature correspondences and gyroscopic data from IMU measurerments to compute the motion hypothesis). By exploiting this 2-point motion parametrization, we proposed two algorithms to remove wrong data associations in the feature matching process for case of a 6DoF motion. We showed that in the case of a monocular camera mounted on a quadrotor vehicle, motion priors from IMU can be used to discard wrong estimations in the framework of a 2 -point-RANSAC based approach. The proposed methods have been evaluated on both synthetic and real data and presented at the last ICRA conference [27].

5.8. Compiling Probabilistic Programs Onto Reconfigurable Logic Using Stochastic Arithmetic

Participants: Emmanuel Mazer, Marvin Faix.

It is of great interest to perform light weight probabilistic inferences for applications such as sensor fusion. Our goal is to design systems to perform these inferences without using a Von Newman machine nor standard floating point arithmetic. By addressing the core of how computations are made, we can explore the tradeoffs between system precision with power consumption and computation time, enabling artificial systems with limited resources, such as mobile and embedded systems, to better operate under uncertainty. Figure 15 illustrates the tool-chain, which starts from the specification of the Bayesian Program in Bayesian programming language , and evaluates it on a reconfigurable device.

This study is part of BAMBI (Bottom-up Approaches to Machines dedicated to Bayesian Inference, www.bambi-fet.eu) : a European collaborative research project relying on the theory of Bayesian inference to understand the natural cognition and aiming at designing bio-inspired computing devices.

A Bayesian machine has probability distributions as inputs and returns a probability distribution as output. It is defined by a joint probability distribution on a set of discrete and finite variables: $P(M \wedge D \wedge L)$. Where M, D and L are themselves conjunctions of variables, for example $D = D_1 \wedge ... \wedge D_k$. We define the soft evidences on the variables D_k as the probability distribution $\widetilde{P}(D_k)$. These soft evidences will be the inputs of the Bayesian machine.

So, given the soft evidences $\widetilde{P}(D_k)$ and the joint distribution $P(M \wedge D \wedge L)$, the machine will fulfil the specification if it computes:

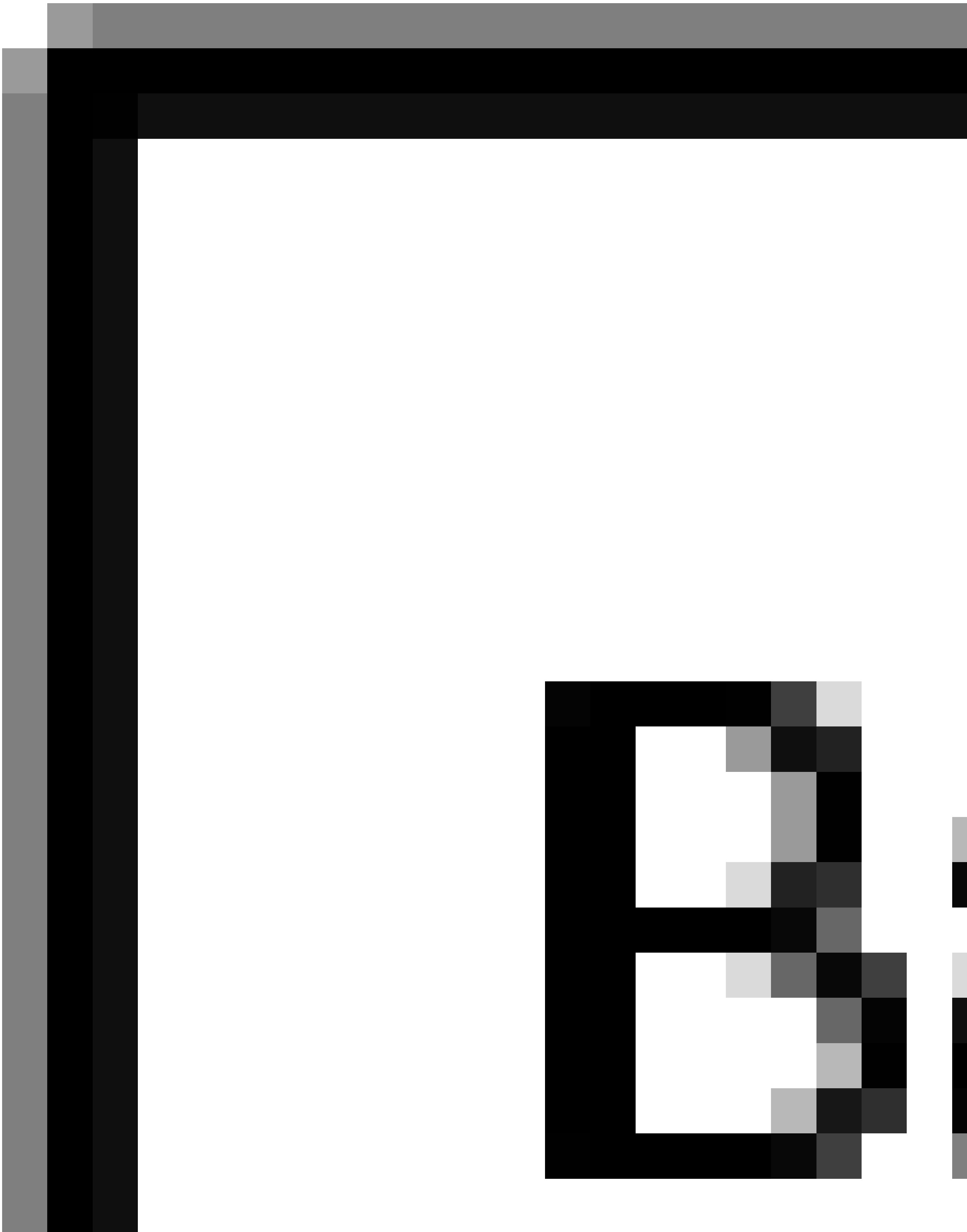
$$P'(M) = \frac{1}{Z}\sum_{D_1}\widetilde{P}(D_1)...\sum_{D_k}\widetilde{P}(D_k)\sum_{L}P(M \wedge D \wedge L) \quad (5)$$

with

$$Z = \sum_{M}(\sum_{D_1}\widetilde{P}(D_1)...\sum_{D_k}\widetilde{P}(D_k)\sum_{L}P(M \wedge D \wedge L)$$

In other words the machine computes a soft inference based on the joint distribution $P(M \wedge D \wedge L)$.

A modified version of the probabilistic language ProBT is used to specify the machine: the joint distribution, the output and the inputs are specified with this language [0]. The next program is an example of a simple specification using the Python bindings of ProBT.



```
#import the ProBT bindings
from pypl import *
#define the variables
dim3 = plIntegerType(0,2)
D1 = plSymbol(D1,dim3)
D2 = plSymbol(D2,dim3)
M= plSymbol(M,dim3)
#define the distribution on M
PM= plProbTable(M,[0.8,0.1,0.1])
#define a conditional distribution on D1
PD1_k_M = plDistributionTable(D1,M)
PD1_k_M.push(plProbTable(D1,[0.5,0.2,0.3]),0)
PD1_k_M.push(plProbTable(D1,[0.5,0.3,0.2]),1)
PD1_k_M.push(plProbTable(D1,[0.4,0.3,0.3]),2)
#define a conditional distribution on D2
PD2_k_M = plDistributionTable(D2,M)
PD2_k_M.push(plProbTable(D2,[0.2,0.6,0.2]),0)
PD2_k_M.push(plProbTable(D2,[0.6,0.3,0.1]),1)
PD2_k_M.push(plProbTable(D2,[0.3,0.6,0.1]),2)
#define the joint distribution
model=plJointDistribution(PM*PD1_k_M*PD2_k_M)
#define the soft evidence variables
model.set_soft_evidence_variables(D1^D2)
#define the output
question=model.ask(M)
```

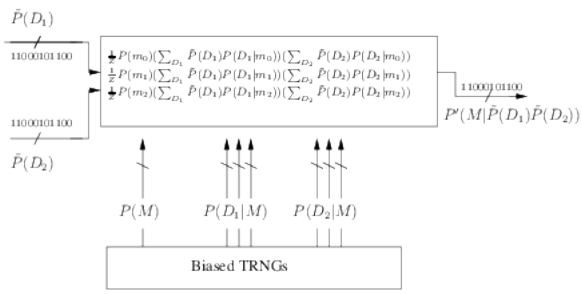


Figure 16. The probabilistic machine corresponding to the given program..

Figure 16 presents the high-level representation of the architecture for the Bayesian Machine. It comprises the main stochastic machine along with the True Random Generators (TRNG), responsible for the generation of the stochastic bit streams for the constants considered in the problem.

[0]A free version of ProBT is available at http://www.probayes.com/fr/Bayesian-Programming-Book/ and the version with soft evidence will be placed on www.bambi-fet.eu before the NIPS conference

The proposed tool-chain is working and accepts any ProBT program with discrete variables as entry. The tool-chain generates a VHDL file which is the description of the stochastic circuit and can be implemented on a FPGA. A Cyclone IV FPGA, from Altera has been targeted as supporting device. A machine has been synthesised to demonstrate the applicability and scalability of the proposed tool-chain. ProBT is also used to compute the exact result using standard arithmetic. This allows to evaluate the results given by FPGA with the synthesised VHDL program.

Figure 17 (right) shows the RTL generated by the synthesis tool, where it is possible to identify the connections between the components, corresponding to the circuit in Figure 17 (left). This circuit was implemented using 6 Logic Elements. The circuit was tested with bit streams integrated over 2^{31} to do the conversion from stochastic to binary.

The proposed tool-chain is working and accepts any ProBT program with discrete variables as entry. The tool-chain generates a VHDL file which is the description of the stochastic circuit and can be implemented on a FPGA. A Cyclone IV FPGA, from Altera has been targeted as supporting device. A machine has been synthesised to demonstrate the applicability and scalability of the proposed tool-chain.

ProBT is also used to compute the exact result using standard arithmetic. This allows to evaluate the results given by FPGA with the synthesised VHDL program.

Figure 17 (right) shows the RTL generated by the synthesis tool, where it is possible to identify the connections between the components, corresponding to the circuit in Figure 17 (left). This circuit was implemented using 6 Logic Elements. The circuit was tested with bit streams integrated over 2^{31} to do the conversion from stochastic to binary.

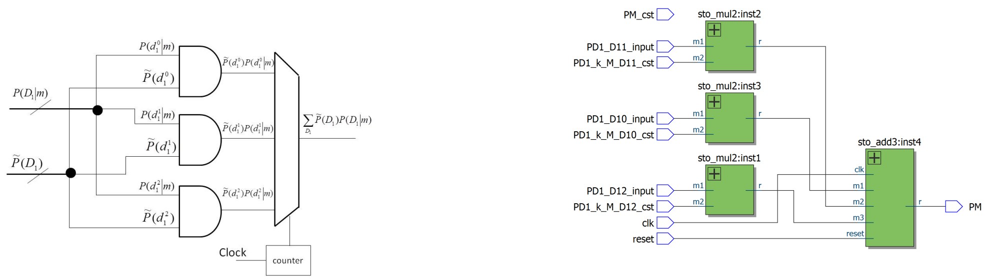


Figure 17. Stochastic circuit computing $\sum_{D1} \widetilde{P}(D_1)P(D_1|m)$ and the corresponding RTL.

We are now focusing on solving the time dilution problem by introducing memory in the architecture. Then we will make an attempt to build a filter with similar ideas by re-fitting the output into the initial joint distribution.

ECUADOR Project-Team

6. New Results

6.1. Resolution of linearised systems

Participants: Olivier Allain [Lemma], Gautier Brèthes, Alain Dervieux, Bruno Koobus, Stephen Wornom.

For Fluid Mechanics as well as for Structural Mechanics, implicit time-advancing is mandatory. It can be applied efficiently if the large systems involved are solved with a good parallel algorithm. In the 90's, a generation of solution algorithms was devised on the basis of Domain Decomposition Methods (DDM). For complex models such as compressible flows, Schwarz DDM were combined with quasi-Newton algorithms like GMRES. For example in the Aironum tool, we use Restrictive Additive Schwarz (RAS, developed by Cai and Farhat). RAS is an ancestor of the widely used class of Newton-Krylov-Schwarz (NKS) algorithms. NKS are, in some versions including RAS, almost scalable *i.e.* their convergence rate is independant of the number of processors. But scalability degrades over a thousand processors. During the ANR ECINADS, coordinated by Ecuador, a Coarse-Grid Deflated RAS was developed. The algorithmic scalability (iteration-wise) holds for all part, except for the coarse grid direct solver, which concerns a much smaller problem. Effective Convergence Scalability (ECS) was confirmed up to 2048 processors. After this level, the asymptotic complexity of the coarse-grid direct solver become predominant and ECS is lost. In other words, with a Coarse-Grid Deflated RAS, the size of the coarse grid problem must be limited in order to enjoy ECS.

In the thesis of Gautier Brèthes, we now study a further step towards super-massive scalability: we use a number of fine and medium grids in order to solve the target large system by a multi-mesh multigrid (MG) algorithm. An important novelty is that the complete FMG technology is applied, with a new stopping criterion avoiding useless cycling [12]. It is well-known that parallel MG is limited to "no-too-coarse" coarse levels due to an excessive ratio between communication and computation. Now our parallel MG can be complemented by the previous version of the solver (deflated RAS) for this no-too-coarse coarse level.

6.2. Algorithmic Differentiation of a CFD code

Participants: Valérie Pascual, Laurent Hascoët, Alain Dervieux.

This activity continues in collaboration with the partners of the FP7 project UMRIDA. The team is assisting Alenia-Aermacchi in the efficient differentiation of its Euler/Navier Stokes UNS3D code in tangent mode, using in particular a differentiable extension of the MPI library.

Inside a collaboration with EDF, Valérie Pascual is also applying Tapenade to produce various adjoint differentiations of the hydrographic code Mascaret.

6.3. Control of approximation errors

Participants: Gautier Brèthes, Eléonore Gauci, Alain Dervieux, Adrien Loseille [Gamma team, Inria-Rocquencourt], Frederic Alauzet [Gamma team, Inria-Rocquencourt], Stephen Wornom, Olivier Allain [Lemma], Anca Belme [University Paris VI].

A study of an interesting combination of Full Multigrid (FMG) and Anisotropic mesh Adaptation (AA) started last year, with the beginning of the thesis of Gautier Brèthes. FMG is one of the (very) few algorithm giving N results by consuming kN floats. Anisotropic adaptation produces approximation errors less than ε with $N = \varepsilon^{-\frac{1}{dim}}$ nodes, this for smooth and non-smooth solution fields. Anisotropic adaptative FMG may produce approximation errors less than ε by consuming $k\varepsilon^{-\frac{1}{dim}}$ floats. Moreover, theory and experiments show that FMG works better when combined with AA. A first AA-FMG platform has been developed. It combines several mesh-adaptation modules developed by Gamma and Distene. It is used for testing new adaptation criteria.

Third-order mesh adaptation was the main topic of last year in error control. The scheme is the ENO finite-volume formulation with quadratic reconstruction. An article describing our results for 2D applications is being written. A 3D version is developed in the Aironum CFD platform. A cooperation with Lemma is also running, with Eléonore Gauci, to apply the scheme to fluid-gas interfaces. Further studies of mesh adaptation for viscous flows are ongoing and an article in collaboration with Gamma3 and University Paris VI (Anca Belme) is being written.

An important novelty in mesh adaptation is the norm-oriented AA method. The method relies on the definition of ad hoc correctors. It has been developed in the academic FMG platform for elliptic problems. Another version is developed by Gamma, in collaboration with Ecuador, for the compressible flow models. The purpose is to devise a composite algorithm in which an approximation error norm can be specified by the user. The introduction of the norm-oriented idea considerably amplifies the impact of adjoint-based AA. The applied mathematician and the engineer now have methods when faced to mesh adaptation for the simulation of a complex PDE system, since they can specify which error norm level they wish, and for which norm. Eléonore Gauci starts a thesis, co-advised by Alain Dervieux and Frédéric Alauzet, on the norm-oriented criteria for CFD and coupled CSM-CFD systems. She also works on a new version of the mesh adaptive CFD demonstrator of Gamma3. This new version improves the resolution of curved features. A cooperation is also starting between Gautier Brèthes and Thierry Coupez (Ecole Centrale de Nantes) on discrete metrics.

These studies are supported by an European FP7 project UMRIDA which deals with the application of AA to approximation error modelling and control, and by ANR project MAIDESC coordinated by Ecuador, which deals with meshes for interfaces, third-order accuracy, meshes for boundary layers, and curved meshes.

6.4. Turbulence models

Participants: Emmanuelle Itam [University Montpellier II], Alain Dervieux, Bruno Koobus, Carine Moussaed [University Montpellier II], Maria-Vittoria Salvetti [University of Pisa], Stephen Wornom, Bruno Sainte-Rose [Lemma].

The purpose of our work in hybrid RANS/LES is to develop new approaches for industrial applications of LES-based analyses. This year, many experiments have validated the dynamic version of our VMS-LES. The quality of simulations is either comparable to non-dynamic, or better. In the applications targetted (aeronautics, hydraulics), the Reynolds number can be as high as several tenth millions, far too high for pure LES models. However, certain regions in the flow can be better predicted with LES than with usual statistical RANS (Reynolds averaged Navier-Stokes) models. These are mainly vortical separated regions as assumed in one of the most popular hybrid model, the hybrid Detached Eddy Simulation model. Here, “hybrid” means that a blending is applied between LES and RANS. An important difference between a real life flow and a wind tunnel or basin is that the turbulence of the flow upstream of each body is not well known. This year, we have started the study of multiple-body flows. A typical case is the interaction between two parallel cylinders, one being in the wake of the other. A recent workshop showed the rather disastrous predictions of LES models. Most hybrid models behave better, mainly for the first cylinder. We are progressively extending and validating our VMS-LES model and our hybrid ones ([11]).

6.5. AD tools infrastructure

Participants: Laurent Hascoët, Paul Hovland [Argonne National Lab. (Illinois, USA)], Sri Hari Krishna Narayanan [Argonne National Lab. (Illinois, USA)].

We have an ongoing partnership with Paul Hovland’s team at Argonne National Lab, formalized by joint participation in the Inria-Illinois joint lab on petascale computing and with travels funded by the Partner University Fund (PUF) of the French embassy in the USA.

In this framework, we worked on the goal of blending our AD tool Tapenade with Argonne's tool OpenAD, buy developing bridges between their internal representations, through a common external representation of analyzed programs. This representation called XAIF is based on XML. We have developed running prototypes of these bridges in both directions, that run on a few examples and that need further development to allow each tool to take advantage of the other's analyses and models. This was supported by two visits of Krishna Narayanan to Inria and one of Laurent Hascoët to Argonne.

We also continued joint development of the Adjoinable-MPI library (AMPI) that provides efficient tangent and adjoint differentiation for MPI-parallel codes, independently of the AD tool used (now AdolC, dco, OpenAD, Tapenade).

We also extracted from Tapenade a standalone kernel (with documented API) for program parsing, analysis, and unparsing, which is not specific to AD and which could be used to develop other source-to-source code transformations. Paul Hovland's team and another Argonne team have shown interest for this library.

6.6. Algorithmic Differentiation and Dynamic Memory

Participants: Laurent Hascoët, Sri Hari Krishna Narayanan [Argonne National Lab. (Illinois, USA)].

In the same framework as in section 6.5 , we made progress in the development of the adjoint AD model for programs that use dynamic memory. Adjoint differentiated code obtained by source transformation (OpenAD, Tapenade...) consists of a forward sweep that essentially copies the original code, and a backward sweep that computes the derivatives, These two sweeps must have the same control-flow shape, only reversed. The allocation and deallocation of dynamic memory inside the forward sweep requires a similar pattern in the backward sweep. However, allocations do not always return the same memory chunk, and therefore all memory addresses must be updated to preserve their consistency in the backward sweep.

This problem can only be solved at run-time. A compile-time analysis simply cannot extract the information needed. Our approach is thus to design a library that encapsulates all calls to memory allocation primitives (`malloc, free`...) in order to register the allocated addresses and to restore consistency of pointers during the backward sweep. This strategy is similar to the one we use for MPI calls, *cf* 6.5 , and is actually an ingredient in our AMPI strategy.

This approach was tested with success on a medium-size industrial application in structural mechanics. For this unsteady simulation the C code allocates and frees memory repeatedly at each time step. The tangent and adjoint differentiated C codes, as produced by Tapenade, have been adapted by hand to run the new model, showing promising performance. Obviously, the next step is to update the Tapenade AD model to automate this hand adaptation.

6.7. Algorithmic Differentiation and Iterative Processes

Participants: Laurent Hascoët, Ala Taftaf.

Adjoint codes naturally propagate partial gradients backwards from the result of the simulation. However, this uses the data flow of the simulation in reverse order, at a cost that increases with the length of the simulation. In the special case of iterative Fixed-Point loops, it is clear that the first iterations operate on a meaningless "initial guess" state vector, and that reversing the corresponding data-flow is wasted effort. An adapted adjoint strategy for the iterative process should consider only the last or the few last iterations. Also the adjoint loop, which is itself a Fixed-Point iteration, must have its own stopping criterion and not merely run as many times as the direct Fixed-Point loop. We selected the strategy proposed by Bruce Christianson [17] and this year we implemented it in Tapenade. This strategy is triggered by differentiation directives that we defined. We tested this strategy with success on the medium-size testcase provided by Queen Mary University for the AboutFlow project.

Ala Taftaf presented her results at the WCCM congress during the Eccomas conference in Barcelona [13], july 21-25. Ala Taftaf did a two-months secondment for her Marie Curie PhD grant, with our partner team of Queen Mary University of London, during which she helped them take advantage of the latest developments in Tapenade and of her developments about Fixed-Point adjoints.

6.8. Multi-Activity specialized Differentiation

Participants: Laurent Hascoët, Ian Hueckelheim [Queen Mary University of London].

Up to this year, Tapenade did not allow for specialization of differentiated routines for different "activity" patterns. If a procedure must be differentiated once with respect to some of its arguments, and once with respect to another subset of arguments, then only one generalized differentiated procedure is created, with respect to the union of all subsets of active arguments. This incurs an efficiency penalty, but avoids a combinatorial explosion of the differentiated code.

However, there are cases where the efficiency penalty is high, and some users want more flexibility. Also the specialized adjoint for Fixed-Point iterations *cf* 6.7 uses two distinct activity patterns for the Fixed-Point loop body, and merging them looses some of the benefits of the approach. We have modified Tapenade to perform activity-specialized differentiation, giving the end-user a complete control through differentiation directives.

The experiments on a non-contrived industrial testcase of the AboutFlow project showed a non-negligible improvement between 5 to 10%. Work is still in progress to incorporate this new functionality into the mainstream distributed Tapenade. Ian Hueckelheim presented these results at the 16$^{\text{th}}$ EuroAD workshop in Jena, Germany, December 8-9.

ESTASYS Exploratory Action

6. New Results

6.1. Highlights of the Year

The Plasma statistical model checker has been made available to other scientists. ESTASYS has open a new branch on verifying the security of complex systems.

6.2. Verification of Heterogeneous Systems

Participants: Axel Legay, Benoît Boyer, Ngo Van-Chan, Jean Quilbeuf.

> This part concerns Tasks 1, 2 and 4 of the action. We characterize and formalize heterogeneous aspects of SoS and then we define efficient monitoring algorithms and representations for their requirements. We then combine the results with Statistical Model Checking (Task 5).

Systems of Systems (SoS) are very large scale systems with particular characteristics. SoS are not directly built from scratch by a single designer or a single team but are obtained as the composition of simpler systems. SoS have strong reliability and dependability requirements, as they aim to provide a service over a long running period. SoS may dynamically modify themselves by connecting to new systems, updating or disconnecting faulty ones, making it impossible to statically know the set of subsystems that are part of the SoS before runtime.

One of the main difficulty arising when developing SoS is the fact that subsystems may have been designed with a different goal in mind. In particular, some subsystems may have their own goal which differs from the global goal of the SoS. Furthermore, each subsystem may be developed in a particular computation model, making it difficult to find a common unifying semantics for the whole SoS. Finally, SoS may exhibit some emergent behaviors that are hardly predictable at design time.

One of the solutions to allow simulation of a SoS is to rely on a common interface for interconnecting the subsystems. The Functional Mockup Interface (FMI) standard is a natural candidate for such an interface. The different components of a SoS developed in different models of computation can be translated to Functional Mockup Units (FMU). Then a so-called master algorithm coordinates the FMUs composing the system. The execution of each FMU is either directly handled by the master algorithm or relies on an external tool for its execution.

Because the subsystems composing a SoS are of heterogeneous nature, it is difficult to find a common semantics model for the whole system. Furthermore, building such a transition system is not tractable due to the complexity of the system. Thus verification through traditional model checking is not possible for SoS. However, since the FMI/FMU framework enables simulation of such systems, the statistical model checking approach can be used.

The DANSE EU project aims to provide a complete tool chain from the modeling to the verification of SoS. At the higher level, the modeling is done in UPDM using the RHAPSODY tool. At the same level, the designer can express requirements over the model using some patterns written in GCSL. The UPDM model can then be translated into a FMI/FMU format that can be simulated by a dedicated tool, named DESYRE. Similarly, the GCSL requirements are transformed into BLTL formulas. Finally, the PLASMA statistical model checker has been integrated with the DESYRE tool chain in order to check the BLTL formulas based on the simulations provided by DESYRE.

6.2.1. Papers:

[45] (W) This report presents some of the results of the first year of Danse, one of the first EU IP projects dedicated to System of Systems. Concretely, we offer a tool chain that allows to specify SoS and SoS requirements at high level, and analyse them using powerful toolsets coming from the formal verification area. At the high level, we use UPDM, the system model provided by the british army as well as a new type of contract based on behavioral patterns. At low level, we rely on a powerful simulation toolset combined with recent advances from the area of statistical model checking. The approach has been applied to a case study developed at EADS Innovation Works.

[51] (W) Exhaustive formal verification for systems of systems (SoS) is impractical and cannot be applied on a large scale. In this paper we propose to use statistical model checking for efficient verification of SoS. We address three relevant aspects for systems of systems: 1) the model of the SoS, which includes stochastic aspects; 2) the formalization of the SoS requirements in the form of contracts; 3) the tool-chain to support statistical model checking for SoS. We adapt the SMC technique for application to heterogeneous SoS. We extend the UPDM/SysML specification language to express the SoS requirements that the implemented strategies over the SoS must satisfy. The requirements are specified with a new contract language specifically designed for SoS, targeting a high-level English-pattern language, but relying on an accurate semantics given by the standard temporal logics. The contracts are verified against the UPDM/SysML specification using the Statistical Model Checker (SMC) PLASMA combined with the simulation engine DESYRE, which integrates heterogeneous behavioral models through the functional mock-up interface (FMI) standard. The tool-chain allows computing an estimation of the satisfiability of the contracts by the SoS. The results help the system architect to trade-off different solutions to guide the evolution of the SoS.

6.3. Formal Models for Variability

Participants: Axel Legay, Rudolf Fahrenberg, Jin Hyun Kim.

> This part of the report is more concerned with task 2. It studies variability aspects in the broad scope. To simplify the study for the first year, we use the concept of software product lines. Later we shall use the results in federation of embedded systems, which is a particular class of Systems of systems.

Variability is ubiquitous in today's systems, be it in the form of configuration options or extensible architectures. By mastering variability, developers can adapt their system to changing requirements without having to develop entirely new applications. Variability is central in the context of SoS whose behaviors depend on interconnected objects. To gain information on managing variability, we have focused on Software Product Lines. Software Product Lines (SPLs) are a popular form of variability-intensive systems. They are families of similar software systems developed together to make economies of scale. SoS can be viewed as examples of product lines with interconnected obbjects. SPL engineering aims to facilitate the development of the members of a family (called *products* or *variants*) by identifying upfront their commonalities and differences. Variability in SPLs is commonly represented in terms of *features*, *i.e.*, units of difference between products that appear natural to stakeholders. Each product of an SPL is therefore defined by its set of features. Hierarchies of features and dependencies between features (*e.g.*, requires, excludes) are typically captured in a *Feature Model* (FM), *i.e.* a tree-like structure that specifies which combinations of features are valid.

6.3.1. Papers:

[15] (C) The model-checking problem for Software Products Lines (SPLs) is harder than for single systems: variability constitutes a new source of complexity that exacerbates the state-explosion problem. Abstraction techniques have successfully alleviated state explosion in single-system models. However, they need to be adapted to SPLs, to take into account the set of variants that produce a counterexample. In this paper, we apply CEGAR (Counterexample-Guided Abstraction Refinement) and we design new forms of abstraction specifically for SPLs. We carry out experiments to

evaluate the efficiency of our new abstractions. The results show that our abstractions, combined with an appropriate refinement strategy, hold the potential to achieve large reductions in verification time, although they sometimes perform worse. We discuss in which cases a given abstraction should be used.

•

[18] (C) In this work, We explore how ideas of statistical testing, based on a usage model (a Markov chain), can be used to extract configurations of interest according to the likelihood of their executions. These executions are gathered in featured transition systems, compact representation of SPL behaviour. We discuss possible scenarios and give a prioritization procedure validated on a web-based learning management software.

6.4. Statistical Model Checking

Participants: Axel Legay, Sean Sedwards, Benoît Boyer, Louis-Marie Traonouez, Kevin Corre.

This section covers Tasks 4 and 5 of the action. It consists in developping Simulation based techniques and efficient statistical algorithms for SoS.

The use of test cases remains the default means of ensuring the correct behaviour of systems in industry, but this technique is limited by the need to hypothesise scenarios that cause interesting behaviour and the fact that a reasonable set of test cases is unlikely to cover all possible eventualities. Static analysis is more thorough and has been successful in debugging very large systems, but its ability to analyse complex dynamical properties is limited. In contrast, model checking is an exhaustive technique that verifies whether a system satisfies a dynamical temporal logic property under all possible scenarios. For nondeterministic and probabilistic systems, numerical model checking quantifies the probability that a system satisfies a property. It can also be used to quantify the expected cost or reward of sets of executions.

Numerical model checking gives precise, accurate and certain results by exhaustively exploring the state space of the model, however the exponential growth of the state space with system size (the ‘state explosion problem) typically limits its applicability to “toy” systems. Symbolic model checking using efficient data structures can make certain very large models tractable. It may also be possible to construct simpler but behaviourally equivalent models using various symmetry reduction techniques, such as partial order reduction, bisimulation and lumping. If a new system is being constructed, it may be possible to guarantee the overall behaviour by verifying the behaviour of its subcomponents and limiting the way they interact. Despite these techniques, however, the size, unpredictability and heterogeneity of real systems usually make numerical techniques infeasible. Moreover, even if a system has been specified not to misbehave, it is nevertheless necessary to check that it meets its specification.

Simulation-based approaches are becoming increasingly tractable due to the availability of high performance parallel hardware and algorithms. In particular, statistical model checking (SMC) combines the simplicity of testing with the formality of numerical model checking. The core idea of SMC is to create multiple independent execution traces of a system and count how many satisfy a property specified in temporal logic. The proportion of satisfying traces is an estimate of the probability that the system satisfies the property. By thus modelling the executions of a system as a Bernoulli random variable, the absolute error of the estimate can be bounded using, for example, a confidence interval or a Chernoff bound. It is also possible to use efficient sequential hypothesis testing, to decide with specified statistical confidence whether the probability of a property is above or below a given threshold. Since SMC requires multiple independent simulations, it may be efficiently divided on parallel computer architectures, such as grids, clusters, clouds and general purpose computing on graphics processors (GPGPU).

Knowing a result with less than 100% confidence is often sufficient in real applications, since the confidence bounds may be made arbitrarily tight. Moreover, a swiftly achieved approximation may prevent a lot of wasted time during model design. For many complex systems, SMC offers the only feasible means of quantifying performance. Historically relevant SMC tools include APMC, YMER and VESTA. Well-established numerical model checkers, such as PRISM and UPPAAL, are now also including SMC engines. Dedicated SMC tools under active development include COSMOS and our own tool PLASMA. Recognising that SMC may be applied to any discrete event trace obtained by stochastic simulation, we have devised PLASMA-lab, a modular library of SMC algorithms that may be used to construct domain-specific SMC tools. PLASMA-lab has become the main vehicle of our ongoing development of SMC algorithms.

Our group is devising cutting edge techniques for SMC. In particular, we are developing new learning algorithms (Sect. 6.4.3), algorithms for nondeterministic systems (Sect. 6.4.1), and algorithms for rare events (Sect. 6.4.2).

6.4.1. Algorithms for Nondeterminism

Nondeterministic models are of fundamental importance in defining complexity and are useful models of concurrency optimisation problems. This latter application is of particular importance in the context of systems constructed from subsystems ("Systems of Systems") that interact in an unpredictable way. Verifying or optimising such systems is problematic for numerical techniques because the state space is typically intractable. Nondeterminism is challenging for simulation-based techniques because, by definition, an executable semantics is not determined.

We have thus begun a line of research to develop SMC algorithms for nondeterministic systems. Our initial focus is Markov decision processes (MDP), however we are in the process of extending our work to various nondeterministc timed automata. Recent attempts to provide approximative algorithms for MDPs either do not address the standard verification problems, consider only a "spurious" subset of the standard problems or contain significant misconceptions and limitations.

In [28], we presented the first complete set of scalable SMC algorithms for MDPs. Our techniques are based on the idea of encoding a history-dependent scheduler as the seed of a pseudo-randomised hash function. Schedulers are thus chosen at random by selecting random seeds. The possibly infinite behaviour of the scheduler is completely encoded in $\mathcal{O}(1)$ memory. We presented simple sampling algorithms to find optimal schedulers and constructed the statistical confidence bounds necessary to find the optima of multiple estimates.

In [34] we devised the notion of "smart sampling" to dramatically improve the performance of the simple algorithms presented in [28]. The basic idea is to use part of the simulation budget to generate a crude estimate of the optimal scheduler and to use this information to better allocate the remaining budget. We successfully applied our algorithms to a number of standard case studies from the literature. We also highlighted the limitations of our approach.

The algorithms in [28], [34] find schedulers that minimise or maximise the probability of a property. In [37] we have adapted our algorithms to minimise or maximise the expected reward of a system. This adaptation is not entirely straightforward because the standard definition of reward properties assumes an exhaustive exploration of the state space of the MDP. We have included an implicit hypothesis test to include this assumption. In other respects optimising rewards is less challenging than optimising probabilities because rewards are effectively based on properties having probability 1. We demonstrate the accuracy of our rewards-based algorithms on standard case studies from the literature.

6.4.2. Rare Events in SMC

Rare properties are often highly relevant to system performance (e.g., bugs and system failure are required to be rare) but pose a problem for statistical model checking because they are difficult to observe. Fortunately, rare event techniques such as *importance sampling* and *importance splitting* may be successfully applied to statistical model checking.

In a previous work [50], we explicitly considered the use of importance sampling in the context of statistical model checking. We presented a simple algorithm that uses the notion of cross-entropy to find the optimal parameters for an importance sampling distribution. In contrast to previous work, our algorithm uses a low dimensional vector of parameters to define this distribution and thus avoids the often intractable explicit representation of a transition matrix. We showed that our parametrisation leads to a unique optimum and can produce many orders of magnitude improvement in simulation efficiency. We demonstrated the efficacy of our methodology by applying it to models from reliability engineering and biochemistry.

Our contribution [49] was the first attempt to use importance splitting with SMC to overcome the Rare Event problem. The basic idea is to decompose a logical property into nested properties whose probabilities are easier to estimate. Importance splitting achieves this by estimating a sequence of conditional probabilities, whose product is the required result. To apply this idea to model checking it is necessary to define a score function based on logical properties, and a set of levels that delimit the conditional probabilities. We described the necessary and desirable properties of score functions and levels. We illustrated how a score function may be derived from a property and gave two importance splitting algorithms: one that uses fixed levels and one that discovers optimal levels adaptively.

6.4.3. SMC with Changes and Simulink

We have proposed a new SMC algorithm for detecting probability changes in dynamic systems. We have adapted CUSUM, an algorithm that can be used to detect changes in signal monitoring. We show that CUSUM can be used to detect when the probability to satisfy a given property drops below some value. This algorithm offers new possibilities to detect, e.g., emergent behaviors in dynamic systems. Our main contributions has been to extend temporal logic with a change-based operator.

All these SMC algorithms are implemented in PLASMA-Lab, and have been recently exported to MATLAB/Simulink – a widely used environment for modeling, simulating and analyzing multidomain dynamic systems – through an integration of MATLAB/Simulink and PLASMA-lab. This integration exploit MATLAB Control, a library allowing to interact with MATLAB from Java. We have developed two different methods to link the two environments. The first method includes a new plugin for PLASMA-lab that allows to load and execute Simulink models within PLASMA-lab, and therefore apply SMC algorithms to these models. The second method proposes an application that can be launched directly within MATLAB and provide the PLASMA-Lab SMC algorithms.

We have submitted a paper [41] that presents the new CUSUM algorithm and the integration between PLASMA-Lab and Simulink. In this paper, we apply these results to a case-study developed with Simulink that models a temperature controller of a pig shed. We show how to use PLASMA-Lab to check SMC requirements, perform parameters optimisation and detect failures in the model using the new CUSUM algorithm.

6.4.4. Papers

[48] (C) Statistical model checking (SMC) offers the potential to decide and quantify dynamical properties of models with intractably large state space, opening up the possibility to verify the performance of complex real-world systems. Rare properties and long simulations pose a challenge to this approach, so here we present a fast and compact statistical model checking platform, PLASMA, that incorporates an efficient simulation engine and uses importance sampling to reduce the number and length of simulations when properties are rare. For increased flexibility and ef-ficiency PLASMA compiles both model and property into bytecode that is executed on an in-built memory-efficient virtual machine.

[47] (C) We present PLASMA-lab, a statistical model checking (SMC) library that provides the functionality to create custom statistical model checkers based on arbitrary discrete event modelling languages. PLASMA-lab is written in Java for maximum cross-platform compatibility and has already been incorporated in various performance-critical software and embedded hardware platforms. Users need only implement a few simple methods in a simulator class to take advantage of our efficient SMC algorithms. PLASMA-lab may be instantiated from the command line or from within other software. We have constructed a graphical user interface (GUI) that exposes the functionality of PLASMA-lab and facilitates its use as a standalone application with multiple 'drop-in' modelling languages. The GUI adds the notion of projects and experiments, and implements a simple, practical means of distributing simulations using remote clients.

[41] (C; submitted) Statistical Model Checking (SMC) is a powerful and widely used approach that consists in extracting global information on the system by monitoring some of its executions. In this paper, we add two new stones to the cathedral of results on SMC, that are 1. a new algorithm to detect emergent behaviors at runtime, and 2. an integration of Plasma Lab, a powerful SMC checker,

as a library of Simulink. Our results are illustrated on a realistic case study.

[26] (C) In this paper, we make use of the notion of a *score function* to improve the granularity of a logical property. We show that such a score function may take advantage of heuristics, so long as it also rigorously respects certain properties. To demonstrate our importance splitting approach we present an optimal adaptive importance splitting algorithm and an heuristic score function. We give experimental results that demonstrate a significant improvement in performance over alternative approaches.

[43] (C; submitted) We introduce feedback-control statistical system checking (FC-SSC), a new approach to statistical model checking that exploits principles of feedback-control for the analysis of cyber-physical systems (CPS). FC-SSC uses stochastic system identification to learn a CPS model, importance sampling to estimate the CPS state, and importance splitting to control the CPS so that the probability that the CPS satisfies a given property can be efficiently inferred. We illustrate the utility of FC-SSC on two example applications, each of which is simple enough to be easily understood, yet complex enough to exhibit all of FC-SCC's features. To the best of our knowledge, FC-SSC is the first statistical system checker to efficiently estimate the probability of rare events in realistic CPS applications or in any complex probabilistic program whose model is either not available, or is infeasible to derive through static-analysis techniques.

6.5. Quantitative Reasoning

Participants: Axel Legay, Rudolf Fahrenberg, Louis-Marie Traonouez.

This part is concerned with Tasks 1 and 2. Mostly, we focus on quantifying properties of interconnected objects such as CPS (SoS and CPS share a lot of commonalities).

Model checking of systems deals with the question whether a given model of a computer system satisfies the properties one might want to require of it. This is a well-established and successful approach to formal verification of safety-critical computer systems.

When the models of the systems contain quantitative information, which is needed to represent the material on which the SoS is running, the model checking problem becomes complicated by the fact that in most cases, quantitative properties of the systems do not need to be satisfied exactly. Indeed, the model or the properties might be subject to measurement error, or probabilistic information might only be an approximation. In this case, it is of little use to know whether or not a model satisfies a specification precisely; what is needed instead is a notion of *satisfaction distance*: a measure which can assess to which extent a quantitative model satisfies a quantitative specification.

In other words, what is needed is a notion of satisfaction which is robust in the sense that small deviations in the model or the specification only lead to small changes in the outcome of the model checking question.

For reasoning about distributed systems or **systems-of-systems**, an important role is played by specification theories. Such systems are often far too complex to reason about, or model-check, as a whole, and additionally they might be composed of a large number of components which are implemented by different vendors. Hence one needs methods for compositional reasoning, which allow to infer properties of a system from properties of its components, and for incremental design, which allow to synthesize and refine specifications in a step-wise manner.

Such specification theories are by now well-established e.g. in the incarnations of interface theories and (disjunctive) modal transition systems. Additionally to defining a formalism for describing and model-checking specifications, they provide notions of refinement of specifications, logical conjunction of specifications, and structural composition and quotient.

When the models and specifications contain quantitative information, all the above notions need to be made robust. One needs to introduce a quantitative version of refinement, and the operations on specifications need to be continuous with respect to refinement distance: compositions of specifications with small refinement distance need themselves to have small refinement distance.

6.5.1. Theory papers:

[33] (J; submitted) There are two fundamentally different approaches to specifying and verifying properties of systems. The logical approach makes use of specifications given as formulae of temporal or modal logics and relies on efficient model checking algorithms; the behavioural approach exploits various equivalence or refinement checking methods, provided the specifications are given in the same formalism as implementations. In this paper we provide translations between the logical formalism of nu-calculus and the behavioural formalism of disjunctive modal transition systems. The translations preserve structural properties of the specification and allow us to perform logical operations on the behavioural specifications as well as behavioural compositions on logical formulae. The unification of both approaches provides additional methods for component-based stepwise design.

[4] (C) This paper studies a difference operator for stochastic systems whose specifications are represented by Abstract Probabilistic Automata (APAs). In the case refinement fails between two specifications, the target of this operator is to produce a specification APA that represents all witness PAs of this failure. Our contribution is an algorithm that allows to approximate the difference of two APAs with arbitrary precision. Our technique relies on new quantitative notions of distances between APAs used to assess convergence of the approximations, as well as on an in-depth inspection of the refinement relation for APAs. The procedure is effective and not more complex to implement than refinement checking.

[21] (C) We provide a framework for compositional and iterative design and verification of systems with quantitative information, such as rewards, time or energy. It is based on disjunctive modal transition systems where we allow actions to bear various types of quantitative information. Throughout the design process the actions can be further refined and the information made more precise. We show how to compute the results of standard operations on the systems, including the quotient (residual), which has not been previously considered for quantitative non-deterministic systems. Our quantitative framework has close connections to the modal nu-calculus and is compositional with respect to general notions of distances between systems and the standard operations.

[35] (J; submitted) We provide a framework for compositional and iterative design and verification of systems with quantitative information, such as rewards, time or energy. It is based on disjunctive modal transition systems where we allow actions to bear various types of quantitative information. Throughout the design process the actions can be further refined and the information made more precise. We show how to compute the results of standard operations on the systems, including the quotient (residual), which has not been previously considered for quantitative non-deterministic systems. Our quantitative framework has close connections to the modal nu-calculus and is compositional with respect to general notions of distances between systems and the standard operations.

[6] (J) This paper proposes a new theory of quantitative specifications. It generalizes the notions of stepwise refinement and compositional design operations from the Boolean to an arbitrary quantitative setting. Using a great number of examples, it is shown that this general approach permits to unify many interesting quantitative approaches to system design.

[7] (J) We present a distance-agnostic approach to quantitative verification. Taking as input an unspecified distance on system traces, or executions, we develop a game-based framework which allows us to define a spectrum of different interesting system distances corresponding to the given trace distance. Thus we extend the classic linear-time–branching-time spectrum to a quantitative setting, parametrized by trace distance. We also prove a general transfer principle which allows us to transfer counterexamples from the qualitative to the quantitative setting, showing that all system distances are mutually topologically inequivalent.

[25] (C) We introduce a new notion of structural refinement, a sound abstraction of logical implication, for the modal nu-calculus. Using new translations between the modal nu-calculus and disjunctive modal transition systems, we show that these two specification formalisms are structurally equivalent.Using our translations, we also transfer the structural operations of composition and quotient from disjunctive modal transition systems to the modal nu-calculus. This shows that the modal nu-calculus supports composition and decomposition of specifications.

6.5.2. Application papers:

[32] (C; submitted) We suggest a method for measuring the degree to which features interact in feature-oriented software development. We argue that our method is practically feasible, easily extendable and useful from a developer's point of view.

[19] (C) Class diagrams are among the most popular modeling languages in industrial use. In a model-driven development process, class diagrams evolve, so it is important to be able to assess differences between revisions, as well as to propagate differences using suitable merge operations. Existing differencing and merging methods are mainly syntactic, concentrating on edit operations applied to model elements, or they are based on sampling: enumerating some examples of instances which characterize the difference between two diagrams. This paper presents the first known (to the best of our knowledge) automatic model merging and differencing operators supported by a formal semantic theory guaranteeing that they are semantically sound. All instances of the merge of a model and its difference with another model are automatically instances of the second model. The differences we synthesize are represented using class diagram notation (not edits, or instances), which allows creation of a simple yet flexible algebra for diffing and merging. It also allows presenting changes comprehensively, in a notation already known to users.

[20] (C) We propose a new similarity measure between texts which, contrary to the current state-of-the-art approaches, takes a global view of the texts to be compared. We have implemented a tool to compute our textual distance and conducted experiments on several corpuses of texts. The experiments show that our methods can reliably identify different global types of texts.

[23] (C) Reliable model transformations are essential for agile modeling. We propose to employ a configurable-semantics approach to develop automatic model transformations which are correct by design and can be integrated smoothly into existing tools and work flows.

[39] (C; submitted) Nowadays, large software systems are mostly built using existing services. These are not always designed to interact, i.e., their public interfaces often present some mismatches. Checking compatibility of service interfaces allows one to avoid erroneous executions when composing the services and ensures correct reuse and interaction. Service compatibility has been intensively studied, in particular for discovery purposes, but most of existing approaches return a Boolean result. In this paper, we present a quantitative approach for measuring the compatibility degree of service interfaces. Our method is generic and flooding-based, and fully automated by a prototype tool.

6.5.3. Surveys:

[22] Modal transition systems provide a behavioral and compositional specification formalism for reactive systems. We survey two extensions of modal transition systems: parametric modal transition systems for specifications with parameters, and weighted modal transition systems for quantitative specifications.

[24] We survey extensions of modal transition systems to specification theories for probabilistic and timed systems.

6.6. Privacy and Security

Participants: Axel Legay, Fabrizio Biondi, Jean Quilbeuf, Thomas Given-Wilson.

6.6.1. Information-Theoretical Quantification of Security Properties

This part of the work was not foreseen at the beginning of the action. It concerns security aspects, and more precisely quantifying privacy of data. This aspect is in fact central for SoS and all our algorithms developed for Tasks 4 and 5 should be adapted to solve a series of problems linked to privacy in interconnected object and dynamical environment. For now, we only studied the foundations.

Information theory provides a powerful quantitative approach to measuring security and privacy properties of systems. By measuring the *information leakage* of a system security properties can be quantified, validated, or falsified. When security concerns are non-binary, information theoretic measures can quantify exactly how much information is leaked. The knowledge of such informations is strategic in the developments of component-based systems.

The quantitative information-theoretical approach to security models the correlation between the secret information of the system and the output that the system produces. Such output can be observed by the attacker, and the attacker tries to infer the value of the secret by combining this information with its knowledge of the system.

Armed with the produced output and the source code of the system, the attacker tries to infer the value of the secret. The quantitative analysis we implement computes with arbitrary precision the number of bits of the secret that the attacker will expectedly infer. This expected number of bits is the information leakage of the system.

The quantitative approach generalizes the qualitative approach and thus provides superior analysis. In particular, a system respects non-interference if and only if its leakage is equal to zero. In practice very few systems respect non-interference, and for those who don't it is imperative to be able to distinguish between the ones leaking a very small amount of bits and the ones leaking a significant amount of bits, since only the latter are considered to pose a security vulnerability to the system.

Since black box security analyzes are immediately invalidated whenever an attacker gains information about the source code of the system, we assume that the attacker has a white box view of the system, meaning that it has access to the system's source code. This approach is also consistent with the fact that many security protocol implementations are in fact open source.

The scope of modern software projects is too large to be analyzed manually. For this reason we provide tools that can support the analyst and locate security vulnerabilities in large codebases and projects. We work with a variety of tools, including commercial software analysis tools being adapted with our techniques, and tools such as QUAIL developed here by our team.

We applied the leakage analysis provided by QUAIL to several case studies. Our case studies (voting protocol and smart grid coordination) have in common that a publicly disclosed information is computed from the secret of every participant in the model. In the voting example, the vote of a given voter is secret, but the number of votes for each candidates is public. Similarly, in the smart grid example, the consumption of one of the houses is secret, but the consumption of a whole quarter can be deduced. Qualitative analyses are either too restrictive or too permissive on these types of systems. For instance, non-interference will reject them as the public information depends on the secret. Declassification approaches will accept them, even if the number of voters or consumers is 2, in which case the secret can be deduced.

The development of better tools for quantitative security builds upon both theoretical developments in information theory, and development of the tools themselves. These often progress in parallel with each supporting the findings of the other, and increasing the demands and understanding upon each other.

6.6.1.1. Papers:

[3] (J; submitted) The quantification of information leakage provides a quantitative evaluation of the security of a system. We propose the usage of Markovian processes to model deterministic and probabilistic systems. By using a methodology generalizing the lattice of information approach we model refined attackers capable to observe the internal behavior of the system, and quantify the information leakage of such systems. We also use our method to obtain an algorithm for the computation of channel capacity from our Markovian models. Finally, we show how to use the method to analyze timed and non-timed attacks on the Onion Routing protocol.

[46] (C) Quantitative security analysis evaluates and compares how effectively a system protects its secret data. We introduce QUAIL, the first tool able to perform an arbitrary-precision quantitative analysis of the security of a system depending on private information. QUAIL builds a Markov Chain model of the system's behavior as observed by an attacker, and computes the correlation between

the system's observable output and the behavior depending on the private information, obtaining the expected amount of bits of the secret that the attacker will infer by observing the system. QUAIL is able to evaluate the safety of randomized protocols depending on secret data, allowing to verify a security protocol's effectiveness. We experiment with a few examples and show that QUAIL's security analysis is more accurate and revealing than results of other tools.

[40] (C; submitted) Quantitative security techniques have been proven effective to measure the security of systems against various types of attackers. However, such techniques are based on computing exponentially large channel matrices or Markov chains, making them impractical for large programs. We propose a different approach based on abstract trace analysis. By analyzing directly sets of execution traces of the program and computing security measures on the results, we are able to scale down the exponential cost of the problem. Also, we are able to appy statistical simulation techniques, allowing us to obtain significant results even without exploring the full space of traces. We have implemented the resulting algorithms in the QUAIL tool. We compare their effectiveness against the state of the art LeakWatch tool on two case studies: privacy of user consumption in smart grid systems and anonymity of voters in different voting schemes.

[12] (C) In an election, it is imperative that the vote of the single voters remain anonymous and undisclosed. Alas, modern anonymity approaches acknowledge that there is an unavoidable leak of anonymity just by publishing data related to the secret, like the election's result. Information theory is applied to quantify this leak and ascertain that it remains below an acceptable threshold. We apply modern quantitative anonymity analysis techniques via the state-of-the-art QUAIL tool to the voting scenario. We consider different voting typologies and establish which are more effective in protecting the voter's privacy. We further demonstrate the effectiveness of the protocols in protecting the privacy of the single voters, deriving an important desirable property of protocols depending on composite secrets.

[13] (C) In recent years, quantitative security techniques have been providing effective measures of the security of a system against an attacker. Such techniques usually assume that the system produces a finite amount of observations based on a finite amount of secret bits and terminates, and the attack is based on these observations. By modeling systems with Markov chains, we are able to measure the effectiveness of attacks on non-terminating systems. Such systems do not necessarily produce a finite amount of output and are not necessarily based on a finite amount of secret bits. We provide characterizations and algorithms to define meaningful measures of security for non-terminating systems, and to compute them when possible. We also study the bounded versions of the problems, and show examples of non-terminating programs and how their effectiveness in protecting their secret can be measured.

EXMO Project-Team

6. New Results

6.1. Highlights of the Year

- Our work on link key extraction and evaluation (§6.3.4) has been published at ECAI 2014.
- Jérôme Euzenat has been elected fellow of the European Coordination Committee for Artificial Intelligence (ECCAI).

6.2. Ontology matching and alignments

We pursue our work on ontology matching and alignment support [8] [10] with contributions to evaluation and alignment semantics.

6.2.1. Evaluation

Participant: Jérôme Euzenat.

Since 2004, we run the Ontology Alignment Evaluation Initiative (OAEI) which organises evaluation campaigns for assessing the degree of achievement of actual ontology matching algorithms [2].

This year, we ran the OAEI 2014 evaluation campaign [15]. We used again our generator for generating new version of benchmarks. The Alignment API was used for manipulating alignments and evaluating results.

A novelty of this year was that data interlinking evaluation was using the SEALS platform and a new query-based evaluation track was created.

The participating systems and evaluation results were presented in the 9th Ontology Matching workshop, held in Riva de Garda, Italy. More information on OAEI can be found at http://oaei.ontologymatching.org/.

6.2.2. Algebras of alignment relations

Participants: Armen Inants [Correspondent], Jérôme Euzenat.

Qualitative calculus is the central concept in qualitative binary constraint satisfaction problems. All formalisms developed so far are homogeneous – they assume a single universe. We had previously shown the advantages of using a homogeneous qualitative calculus for expressing ontology alignment relations between concepts. We tackle the problem of combining two or more calculi over disjoint universes into a single calculus. The problem is important, because in the ontology matching domain we deal with various kinds of ontological entities: concepts, individuals, properties. We define a new formalism called a heterogeneous qualitative calculus, based on an algebraic construct called Schröder category. A Schröder category is to binary relations over heterogeneous universes what a relation algebra is to homogeneous ones. We establish the connection between homogeneous and heterogeneous qualitative calculi by defining two mutually inverse transition operators. We provide an algorithm for combining two homogeneous calculi with different universes into a single calculus.

This work has vocation to support developments of the Alignment API towards relation algebras. It is part of the PhD of Armen Inants.

6.3. Data interlinking

The web of data uses semantic web technologies to publish data on the web in such a way that they can be interpreted and connected together. It is thus important to be able to establish links between these data, both for the web of data and for the semantic web that it contributes to feed. We consider this problem from different perspectives.

6.3.1. Interlinking cross-lingual RDF data sets

Participants: Tatiana Lesnikova [Correspondent], Jérôme David, Jérôme Euzenat.

RDF data sets are being published with labels that may be expressed in different languages. Even systems based on graph structure, ultimately rely on anchors based on language fragments. In this context, data interlinking requires specific approaches in order to tackle cross-lingualism. We proposed a general framework for interlinking RDF data in different languages and implemented two approaches: one approach is based on machine translation, the other one is based taking advantage of multilingual refernces, such as BabelNet. We evaluated variation of theses two settings on English (DBPedia) and Chinese (XLore) datasets. Both approaches demonstrated promising results [20]. We will conduct more experiments including other language pairs and larger corpus.

This work is part of the PhD of Tatiana Lesnikova developed in the LINDICLE project (§7.1.2).

6.3.2. Interactive learning of interlinking patterns

Participants: Zhengjie Fan [Correspondent], Jérôme Euzenat.

We proposed an interlinking method which, from class correspondences between data source ontologies, uses k-means or k-medoids clustering to produce property correspondences. It then generates a first interlinking pattern which is a combination of a link key and similarity measures. Such patterns can be transformed into a SILK script for generating an initial link set. A sample of these links are assessed by users as either correct or incorrect. These are taken as positive and negative example by an extension of the disjunctive version space method to find an interlinking pattern, that can justify correct links and incorrect links. Experiments show that, with only 1% of sample links, this method reaches a F-measure over 96%. The F-measure quickly converges, being improved by nearly 10% than other comparable approaches [19].

This work is part of the PhD of Zhengjie Fan [4], co-supervised with François Scharffe (LIRMM), and developed in the DATALIFT project (§7.1.1).

6.3.3. An iterative import-by-query approach to data interlinking

Participants: Manuel Atencia Arcas [Correspondent], Mustafa Al-Bakri, Steffen Lalande, Marie-Christine Rousset.

We modelled the problem of data interlinking as a reasoning problem on possibly decentralised data. We described an import-by-query algorithm that alternates steps of sub-query rewriting and of tailored querying of data sources. It only imports data as specific as possible for inferring or contradicting target sameAs assertions. Experiments conducted on a real-world dataset have demonstrated in practice the feasibility and usefulness of this approach for data interlinking and disambiguation purposes.

This work is part of the PhD thesis of Mustafa Al-Bakri, co-supervised by Manuel Atencia and Marie-Christine Rousset, developed in the QUALINCA project.

6.3.4. Link key extraction

Participants: Jérôme David [Correspondent], Manuel Atencia Arcas, Jérôme Euzenat.

Ontologies do not necessarily come with key descriptions, and never with link key assertions. Keys can be extracted from data by assuming that keys holding for specific data sets, may hold universally. We have extended such a classical key extraction technique for extracting weak link keys. We designed an algorithm to generate first a small set of candidate link keys and described this approach in the framework of formal context analysis [13]. Depending on whether some of the, valid or invalid, links are known, we defined supervised and non supervised measures for selecting the appropriate link keys. The supervised measures approximate precision and recall on a sample, while the non supervised measures are the ratio of pairs of entities a link key covers (coverage), and the ratio of entities from the same data set it identifies (discrimination). We have experimented these techniques, showing the accuracy and robustness of both approaches [12].

This work has been developed partly in the LINDICLE project (§7.1.2).

6.4. Dynamic aspects of networks of ontologies

Huge quantities of data described by ontologies and linked together are made available. These are generated in an independent manner by autonomous providers such as individuals or companies. They are heterogeneous and their joint exploitation requires connecting them, ending up as a mesh of reticulated knowledge.

However, data and knowledge have to evolve facing changes in what they represent, changes in the context in which they are used and connections to new data and knowledge sources. As their production and exchange are growing larger and more connected, their evolution is not anymore compatible with manual curation and maintenance. We work towards their continuous evolution as it is critical to their sustainability.

Two different approaches are currently explored.

6.4.1. Evolution of ontology networks and linked data

Participants: Adam Sanchez Ayte [Correspondent], Jérôme David, Jérôme Euzenat.

As link keys are obtained by statistical analysis of datasets (§6.3.4), they are both data-dependent and computation-intensive. Therefore, their recalculation should be avoided if possible. We are developing methods to analyse if changes performed in the data, necessarily require link key recomputation.

To reach this goal, we are developing an approach considering datasets as logical theories. In this context, changes that affect a link key are meta-logical operations. We adopt the framework of belief revision to define postulates that evolution operators must satisfy.

This work is part of the PhD thesis of Adam Sanchez Ayte developed in the LINDICLE project (§7.1.2).

6.4.2. Cultural alignment repair

Participant: Jérôme Euzenat [Correspondent].

Alignments between ontologies may be established through agents holding such ontologies attempting at communicating and taking appropriate action when communication fails. This approach, that we call cultural repair, has the advantage of not assuming that everything should be set correctly before trying to communicate and of being able to overcome failures. We tested this approach on alignment repair, i.e., the improvement of incorrect alignments. For that purpose, we performed a series of experiments in which agents react to mistakes in alignments. Agents only know about their ontologies and alignments with others and they act in a fully decentralised way. We showed that cultural repair is able to converge towards successful communication through improving the objective correctness of alignments. The obtained results are on par with a baseline of state-of-the-art alignment repair algorithms [7] [17]

The benchmarks, results and software are available at http://lazylav.gforge.inria.fr.

FLOWERS Project-Team

6. New Results

6.1. Highlights of the Year

PY. Oudeyer and M. Lopes, together with J. Gottlieb (Univ. Columbia, NY) organized the first International Symposium on Neurocuriosity symposium on Information Seeking, Curiosity and Attention, pioneering a gathering of world experts on curiosity from developmental psychology, neuroscience, ethology and computational modelling (see https://openlab-flowers.inria.fr/t/first-interdisciplinary-symposium-on-information-seeking-curiosity-and-attention/21). This was achieved in the context of associated team Neurocuriosity with the cognitive neuroscience lab of J. Gottlieb at Univ. Columbia, NY, US. The first results investigating predictions of theoretical formal models of curiosity on human exploration were also published [25].

O. Mangin obtained the Best thesis poster from Bordeaux doctoral school of mathematics and computer science, for his PhD thesis "The Emergence of Multimodal Concepts: From Perceptual Motion Primitives to Grounded Acoustic Words" [24].

The team, in collaboration with Inaki Iturrate and Luis Montesano, published major results on calibration-free brain-computer interface methods, where incremental machine learning algorithms are used to remove the phase of calibration for an important family of use contexts [44] [45].

In october 2014, the team announced the release of a new version of the Poppy Project platform, dedicated in particular to the use of tools for creating and programming interactive robots in Education and Art. This platform, which is a result of research on the role of morphology in skill acquisition within ERC project Explorers, was selected as finalist for the Global Fab Awards 2014 (https://www.fab10.org/en/awards) which select the best worldwide projects in the Makers ecosystem. It was also presented in major international press and media (https://www.poppy-project.org/in-the-press/), in multiple hackatons and demos, in particular at the major international conference LeWeb (https://www.poppy-project.org/social-life/, and its video on the web was seen 125k times. Poppy Project was presented at Elysée, during a French Tech event, to François Hollande (http://www.inria.fr/centre/bordeaux/actualites/poppy-le-robot-humanoide-a-l-elysee), and in Bordeaux to Axelle Lemaire. Web site: http://www.poppy-project.org

The Flowers team made major achievements in diffusing science and technology towards the general public. The team developped the IniRobot pedagogical kit, for the discovery of computer science and robotics in primary schools. The kit was first developped and evaluated in schools, in collaboration with a group of teachers, and then began to be largely disseminated and used in september 2014 to schools in Talence, Bordeaux, Lormont, and Lille. A dedicated web site has been created, allowing all users and contributors to share their experiences with the kit: https://dm1r.inria.fr/c/kits-pedagogiques/inirobot. PY. Oudeyer was invited to give a TedX talk (https://www.youtube.com/watch?v=AP8i435ztwE, video viewed by more than 9000 people), and was interviewed and invited to talk about our research on major media channels (e.g. Le Monde, Les Echos, France Inter, see http://www.pyoudeyer.com/press/).

6.2. Robotic And Computational Models Of Human Development

6.2.1. Computational Models Of Information-Seeking, Curiosity And Attention

Participants: Manuel Lopes, Pierre-Yves Oudeyer [correspondant], Jacqueline Gottlieb, Adrien Baranes, Pierre Rouanet, Brice Miard, Jonathan Grizou.

6.2.1.1. The effects of task difficulty, novelty and the size of the search space on intrinsically motivated exploration

Devising efficient strategies for exploration in large open-ended spaces is one of the most difficult computational problems of intelligent organisms. Because the available rewards are ambiguous or unknown during the exploratory phase, subjects must act in intrinsically motivated fashion. However, a vast majority of behavioral and neural studies to date have focused on decision making in reward-based tasks, and the rules guiding intrinsically motivated exploration remain largely unknown. To examine this question we developed a paradigm for systematically testing the choices of human observers in a free play context. Adult subjects played a series of short computer games of variable difficulty, and freely choose which game they wished to sample without external guidance or physical rewards. Subjects performed the task in three distinct conditions where they sampled from a small or a large choice set (7 vs. 64 possible levels of difficulty), and where they did or did not have the possibility to sample new games at a constant level of difficulty. We show that despite the absence of external constraints, the subjects spontaneously adopted a structured exploration strategy whereby they (1) started with easier games and progressed to more difficult games, (2) sampled the entire choice set including extremely difficult games that could not be learnt, (3) repeated moderately and high difficulty games much more frequently than was predicted by chance, and (4) had higher repetition rates and chose higher speeds if they could generate new sequences at a constant level of difficulty. The results suggest that intrinsically motivated exploration is shaped by several factors including task difficulty, novelty and the size of the choice set, and these come into play to serve two internal goals—maximize the subjects' knowledge of the available tasks (exploring the limits of the task set), and maximize their competence (performance and skills) across the task set. This was published in [25].

6.2.1.2. A new experimental setup to study the structure of curiosity-driven exploration in humans

We started evaluating several games that test how humans explore a space of motor tasks of different complexities. Our objective is to observe what exploratory behaviors do people use when learning a new skill. The main hypothesis we are testing is that skills that provide a larger learning progress will be favored and so we will see a progression from the simpler to the more complex skills. Surely there are individual differences and the causes and impact of those differents is a very important research topic. The Abstract Games we created allows us to create several dimensions of complexity for the games. In this task, there are several abstract forms that appear in the screen and the user is able to control them using its own body (tracked using a Kinect sensor), see Fig. 2. The relation between the degrees of freedom and the forms/colors/sizes of the shapes is arbitrary and the user must explore its body to be able to control its behavior. This was published in [58].

6.3. Life-Long Robot Learning And Development Of Motor And Social Skills

6.3.1. Exploration and learning of sensorimotor policies

6.3.1.1. Non-linear regression algorithms for motor skill acquisition: a comparison

Participants: Thibaut Munzer [correspondant], Freek Stulp, Olivier Sigaud.

Endowing robots with the capability to learn is an important goal for the robotics research community. One part of this research is focused on learning skills, where usually two learning paradigms are used sequentially. First, a robot learns a motor primitive by demonstration (or imitation). Then, it improves this motor primitive with respect to some externally defined criterion. We realized a study on how the representation used in the demonstration learning step can influence the performance of the policy improvement step. We provide a conceptual survey of different demonstration learning algorithms and perform an empirical comparison of their performance when combined with a subsequent policy improvement step. These study have been published at the JFPDA conference [61].

During this work, we have discovered that many (batch) regression algorithms (amongst others, locally weighted (projection) regression, Gaussian mixture regression, radial basis function networks, and Gaussian process regression) use only one of two underlying model representations to represent a function: a weighted sum of basis function, or a mixture of linear models. Furthermore, we show that the former is a special case of the latter. This insights provides a deep understanding of the relationship between these algorithms, that, despite being derived from very different principles, use a function representation that can be captured within one unified model. A review article on this topic has been submitted to Neural Networks.

6.3.1.2. Simultaneous On-line Discovery and Improvement of Robotic Skill Options

Participants: Freek Stulp [correspondant], Laura Herlant, Antoine Hoarau, Gennaro Raiola.

The regularity of everyday tasks enables us to reuse existing solutions for task variations. For instance, most door-handles require the same basic skill (reach, grasp, turn, pull), but small adaptations of the basic skill are required to adapt to the variations that exist (e.g. levers vs. knobs). In a joint project with Laura Herlant of Carnegie Mellon University, we developed the algorithm "Simultaneous On-line Discovery and Improvement of Robotic Skills" (SODIRS) that is able to autonomously discover and optimize skill options for such task variations. We formalize the problem in a reinforcement learning context, and use the PI^{BB} algorithm to continually optimize skills with respect to a cost function. SODIRS discovers new subskills, or "skill options", by clustering the costs of trials, and determining whether perceptual features are able to predict which cluster a trial will belong to. This enables SODIRS to build a decision tree, in which the leaves contain skill options for task variations. We demonstrate SODIRS' performance in simulation, as well as on a Meka humanoid robot performing the ball-in-cup task. This work has led to a publication at IROS [64].

6.3.1.3. Simultaneous On-line Discovery and Improvement of Robotic Skill Options

Participants: Freek Stulp [correspondant], Nicolas Alberto Torres, Michael Mistry.

Freek Stulp supervised the Master's thesis project of Nicolas Torres Alberto from the Telecom Physique Strasbourg, which led to a publication at Humanoids'14 [65]. The project focused on improving autonomy in learning inverse dynamics models for computed torque control. In computed torque control, robot dynamics are predicted by dynamic models. This enables more compliant control, as the gains of the feedback term can be lowered, because the task of compensating for robot dynamics is delegated from the feedback to the feedforward term. Previous work has shown that Gaussian process regression is an effective method for learning computed torque control, by setting the feedforward torques to the mean of the Gaussian process. We extend this work by also exploiting the variance predicted by the Gaussian process, by lowering the gains if the variance is low. This enables an automatic adaptation of the gains to the uncertainty in the computed torque model, and leads to more compliant low-gain control as the robot learns more accurate models over time. On a simulated 7-DOF robot manipulator, we demonstrate how accurate tracking is achieved, despite the gains being lowered over time. This is a first step towards life-long learning robots, that continuously and autonomously adapt their control parameters (feedforward *and* feedback) over extended periods of time.

6.3.1.4. Learning manipulation of flexible tools

Participants: Clément Moulin-Frier [correspondant], Marie-Morgane Paumard, Pierre Rouanet.

Clément Moulin-Frier and Pierre Rouanet supervised the internship of Marie-Morgane Paumard from the *Ecole Normale Supérieure de Cachan*, at the Bachelor level. The internship has been realized from May to August 2014. Her report is entitled *Learning the manipulation of flexible tools in developmental robotics: a fishing robot* and is available at this address: https://flowers.inria.fr/clement_mf/files/Paumard_RapportDeStage.pdf.

Learning how to manipulate flexible tools is an harsh issue in robotics, since there is generally no analytical model of the system dynamics available. Learning algorithms are therefore a pivotal tool to control such systems. Marie-Morgane conceived an experiment on the manipulation of a fishing rod by a 2-arm robot equipped with a movement generation and perceptual systems. She studied how an optimization algorithm allows the robot to reach particular position of the hook on the floor. Then, she analyzed the distribution of effects (i.e. final fishhook position) in different contexts as well as optimization performances for particular goals.

6.3.1.5. Learning how to reach various goals by autonomous interaction with the environment: unification and comparison of exploration strategies

Participants: Clément Moulin-Frier [correspondant], Pierre-Yves Oudeyer.

In the field of developmental robotics, we are particularly interested in the exploration strategies which can drive an agent to learn how to reach a wide variety of goals. We unified and compared such strategies, recently shown to be efficient to learn complex non-linear redundant sensorimotor mappings. They combine two main principles. The first one concerns the space in which the learning agent chooses points to explore (motor space vs. goal space). Previous works (Rolf et al., 2010; Baranes and Oudeyer, 2012) have shown that learning redundant inverse models could be achieved more efficiently if exploration was driven by goal babbling, triggering reaching, rather than direct motor babbling. Goal babbling is especially efficient to learn highly redundant mappings (e.g the inverse kinematics of a arm). At each time step, the agent chooses a goal in a goal space (e.g uniformly), uses the current knowledge of an inverse model to infer a motor command to reach that goal, observes the corresponding consequence and updates its inverse model according to this new experience. This exploration strategy allows the agent to cover the goal space more efficiently, avoiding to waste time in redundant parts of the sensorimotor space (e.g executing many motor commands that actually reach the same goal). The second principle comes from the field of active learning, where exploration strategies are conceived as an optimization process. Samples in the input space (i.e motor space) are collected in order to minimize a given property of the learning process, e.g the uncertainty (Cohn et al., 1996) or the prediction error (Thrun, 1995) of the model. This allows the agent to focus on parts of the sensorimotor space in which exploration is supposed to improve the quality of the model. In [59], we have shown how an integrating probabilistic framework allows to model several recent algorithmic architectures for exploration based on these two principles, and compare the efficiency of various exploration strategies to learn how to uniformly cover a goal space. This was published in [59].

6.3.1.6. Reusing Motor Commands to Learn Object Interaction

Participants: Fabien Benureau [correspondant], Pierre-Yves Oudeyer.

We have proposed the Reuse algorithm, that exploit data produced during the exploration of an first environment to efficiently bootstrap the exploration of second, different but related environment. The effect of the Reuse algorithm is to produce a high diversity of effects early during exploration. The algorithm only constrains the environments to share the same motor space, and makes no assumptions about learning algorithms or sensory modalities. We have illustrated our algorithm on a 6-joints robotic arm interacting with a virtual object, and showed that our algorithm is robust to dissimilar environments, and significantly improves the early exploration of similar ones. This was published in [34].

6.3.1.7. Socially Guided Intrinsic Motivation for Robot Learning of Motor Skills

Participants: Mai Nguyen [correspondant], Pierre-Yves Oudeyer.

We have presented a technical approach to robot learning of motor skills which combines active intrinsically motivated learning with imitation learning. Our architecture, called SGIM-D, allows efficient learning of high-dimensional continuous sensorimotor inverse models in robots, and in particular learns distributions of parameterised motor policies that solve a corresponding distribution of parameterised goals/tasks. This is made possible by the technical integration of imitation learning techniques within an algorithm for learning inverse models that relies on active goal babbling. In an experiment where a robot arm has to learn to use a flexible fishing line , we have illustrated that SGIM-D efficiently combines the advantages of social learning and intrinsic motivation and benefits from human demonstration properties to learn how to produce varied outcomes in the environment, while developing more precise control policies in large spaces. This was published in [28].

6.3.1.8. A social learning formalism for learners trying to figure out what a teacher wants them to do

Participants: Thomas Cederborg [correspondant], Pierre-Yves Oudeyer.

We have elaborated a theoretical foundation for approaching the problem of how a learner can infer what a teacher wants it to do through strongly ambiguous interaction or observation. This groups the interpretation of a broad range of information sources under the same theoretical framework. A teacher's motion demonstration, eye gaze during a reproduction attempt, pushes of good/bad buttons and speech comment are all treated as specific instances of the same general class of information sources. These sources all provide (partially and ambiguously) information about what the teacher wants the learner to do, and all need to be interpreted concurrently. We introduce a formalism to address this challenge, which allows us to consider various strands of previous research as different related facets of a single generalized problem. In turn, this allows us to identify important new avenues for research. To sketch these new directions, several learning setups were introduced, and algorithmic structures are introduced to illustrate some of the practical problems that must be overcome. This was published in [26].

6.3.2. Task learning from social guidance

6.3.2.1. Inverse Reinforcement Learning in Relational Domains

Participants: Thibaut Munzer [correspondant], Bilal Piot, Mathieu Geist, Olivier Pietquin, Manuel Lopes.

We introduced a first approach to the Inverse Reinforcement Learning (IRL) problem in relational domains. IRL has been used to recover a more compact representation of the expert policy leading to better generalize among different contexts. Relational learning allows one to represent problems with a varying number of objects (potentially infinite), thus providing more generalizable representations of problems and skills. We show how these different formalisms can be combined by modifying an IRL algorithm (Cascaded Supervised IRL) such that it handles relational domains. Our results indicate that we can recover rewards from expert data using only partial knowledge about the dynamics of the environment. We evaluate our algorithm in several tasks and study the impact of several experimental conditions such as: the number of demonstrations, knowledge about the dynamics, transfer among varying dimensions of a problem, and changing dynamics.

6.4. Autonomous And Social Perceptual Learning

6.4.1. Unsupervised and online non-stationary obstacle discovery and modelling using a laser range finder

Participants: Guillaume Duceux, David Filliat [correspondant].

Recognizing objects is an important capability for assistance robots, but most methods rely on vision and a heavy training procedures to be able to recognize some objects. Using laser range finders has shown its efficiency to perform mapping and navigation for mobile robots. However, most of existing methods assume a mostly static world and filter away dynamic aspects while those dynamic aspects are often caused by non-stationary objects which may be important for the robot task. We propose an approach that makes it possible to detect, learn and recognize these objects through a multi-view model, using only a planar laser range finder. We show using a supervised approach that despite the limited information provided by the sensor, it is possible to recognize efficiently up to 22 different object, with a low computing cost while taking advantage of the large field of view of the sensor. We also propose an online, incremental and unsupervised approach that make it possible to continuously discover and learn all kind of dynamic elements encountered by the robot including people and objects. These results have been published at the IROS conference [40].

6.4.2. Task oriented representations by discriminative modulation of a generative learning method

Participants: Mathieu Lefort, Alexander Gepperth [correspondant].

PROPRE (which stands for PROjection - PREdiction) is a generic and modular unsupervised neural learning paradigm that extracts meaningful concepts of multiple data flows based on predictability across stimuli. It consists on the combination of three modules. First, a topological projection of each data flow on a self-organizing map. Second, a decentralized prediction of each projection activity from each other map activities. Third, a predictability measure that quantifies the prediction error. This measure is used to modulate the projection learning so that to favor the mapping of predictable stimuli across data flows. This model was applied to the visual supervised classification of the pedestrian orientation. The modulation of the visual representation learning by the predictability measure (quantifying the ability to detect the orientation of the pedestrian) improves significantly classification performances of the system independently of the predictability measure used [55]. Moreover, PROPRE provides a combination of interesting functional properties, such as online and incremental learning [56].

6.4.3. Learning of multimodal representations based on the self-evaluation of their predictability power

Participants: Mathieu Lefort, Thomas Kopinski, Alexander Gepperth [correspondant].

PROPRE paradigm (see section 6.4.2) was also applied to the classification of gestures caught from two time-of-flight (ToF) cameras. In this context, the predictability measure acts as a self-evaluation module that biases the learned representations towards stimuli correlated across modalities, i.e. related to the ability of one camera to predict the other one. We show in [57] that this unsupervised multimodal representations learning improves the gesture recognition performance, compared to isolated camera representations learning, even not as much as supervised one.

6.4.4. Resource-efficient online learning of classification and regression tasks

Participants: Mathieu Lefort, Thomas Kopinski, Thomas Hecht, Alexander Gepperth [correspondant].

This activity investigates the coupling of generative and discriminative learning (SOM and regression) to achieve incremental learning that stays resource-efficient when the number of input and output dimensions is high. On the one hand, we apply this technique to sensory classification problems where input dimensionalities can exceed 10000 in the presence of multiple categories. On the other hand, we target the learning of forward and inverse regression models for robotics, possibly combining proproceptive with sensory information which again leads to high data dimensionality. A special kind of regression task we consider in this context is optimal integration of sensory information, where the most likely underlying value must be inferred from several noisy sensor readings. In contrast to popular approaches like XCF or LWPR, our approach achieves efficiency by avoiding a precise partitioning of the input space, relying on a dimensionality-reduced topological projection of the input space instead. While this achives slightly inferior results on standard benchmarks, we can treat high-dimensional incremental learning problems that are inaccessible to other algorithms, and especially to LWPR. This activity has resulted in two submissions to ESANN 2015 and one to IEEE Transactions on Autonomous Mental Development.

6.4.5. Indoor semantic mapping on a mobile robot

Participants: Louis-Charles Caron [correspondant], Alexander Gepperth, David Filliat.

Semantic mapping is the act of storing high-level information in a persistent map of the environment. The semantic information considered here is the identity of objects encountered by a mobile robot in an indoor environment [35]. The robot runs a SLAM algorithm and builds a map using a laser range finder. The semantic information is collected by analysing the point cloud provided by an RGB-D camera mounted on the robot. The choice of features used to describe the objects, the type of fusion and the recognition algorithm influence the overall capacity of the algorithm. Shape features perform very well, but are blind to changes in color. The fusion of different types of features can reduce the recognition rates on some objects but increases the overall figure. This increase is more significant as the number of objects to recognitze gets larger [36]. After running the object recognition algorithm, the identity of the objects is stored alongside the map. The stored information influences future recognition attempts on objects that were already seen by the robot to improve the recognition process. A 3-d map along with a snapshot and the identity of each object seen is displayed to a user.

6.5. Robot Design And Morphological Computation

6.5.1. Rapid morphological exploration with the Poppy humanoid platform.

Participants: Matthieu Lapeyre [correspondant], Steve N'Guyen, Alexandre Le Falher, Pierre-Yves Oudeyer.

In the paper [53], we discuss the motivation and challenges raised by the desire to consider the morphology as an experimental variable on real robotic platforms as well as allowing reproducibility and diffusion of research results in the scientific community. In this context, we present an alternative design and production methodology that we have applied to the conception of Poppy humanoid, the first complete 3D printed open-source and open-hardware humanoid robot. Robust and accessible, it allows exploring quickly and easily the fabrication, the programming and the experimentation of various robotic morphologies. Both hardware and software are open-source, and a web platform allows interdisciplinary contributions, sharing and collaborations. Finally we conduct an experiment to explore the impact of four different foot morphologies on the robot's dynamic when it makes a footstep. We show that such experimentation can easily be achieved and shared in couple of days at almost no cost.

6.6. Educational Technologies

6.6.1. KidLearn

Participants: Manuel Lopes [correspondant], Pierre-Yves Oudeyer, Didier Roy, Benjamin Clement.

Kidlearn is a research project studying how machine learning can be applied to intelligent tutoring systems. It aims at developing methodologies and software which adaptively personalize sequences of learning activities to the particularities of each individual student. Our systems aim at proposing to the student the right activity at the right time, maximizing concurrently his learning progress and its motivation. In addition to contributing to the efficiency of learning and motivation, the approach is also made to reduce the time needed to design ITS systems.

Intelligent Tutoring System (ITS) are computer environments designed to guide students in their learning. Through the proposal of different activities, it provides teaching experience, guidance and feedback to improve learning. The FLOWERS team has developed several computational models of artificial curiosity and intrinsic motivation based on research on psychology that might have a great impact for ITS. Results showed that activities with intermediate levels of complexity, neither too easy nor too difficult but just a little more difficult that the current level, provide better teaching experiences. The system is based on the combination of three approaches. First, it leverages Flowers team's recent models of computational models of artificial curiosity and intrinsic motivation based on research in psychology and neuroscience. Second, it uses state-of-the-art Multi-Arm Bandit (MAB) techniques to efficiently manage the exploration/exploitation challenge of this optimization process. Third, it leverages expert knowledge to constrain and bootstrap initial exploration of the MAB, while requiring only coarse guidance information of the expert and allowing the system to deal with didactic gaps in its knowledge. In 2014, we have run a second pilot experiment in elementary schools of Région Aquitaine, where 7-8 year old kids could learn elements of mathematics thanks to an educational software that presented the right exercises at the right time to maximize learning progress. [69], [37], [38], [39]

6.6.2. Education and the Poppy project

Participants: Matthieu Lapeyre [correspondant], Pierre-Yves Oudeyer, Didier Roy.

The Poppy platform was initially designed for research purposes and even more specifically for studying biped locomotion and human-robot interaction. However, it has been designed with open science goals in mind, both to share our research and create tools for researchers. As we are convinced of the need for multidisciplinary contributions in order to improve the state of the art in the robotics field, we decided right from the beginning to use and create modern and easy-to-use tools. This choice has strongly affected the way we designed our platform. Indeed, being simple to use, easily reproducible and hackable, modular, 3D printable and as plug 'n play as possible lead to the development of hardware (Poppy) and software (pypot) tools that can be also used by non-expert people.

Thus Poppy meets a growing societal need: education and training in technologies combining computer science, electronics and mechanics, as well as a training tool for the emergent revolutionary 3D printing process. Since October 2013 (open source release), we have been contacted by several Fablabs, universities, engineering schools and even high schools. We have had the opportunity to meet with educational teams and it appears they are looking for new motivational tools for group projects.

In this context, the Poppy platform appears well suited. Indeed, it integrates advanced and yet easily accessible techniques (3D printing, Arduino, Python) in an embodiment that motivates students and the wider public. With its openness, design and rather low-cost, Poppy is highly hackable and provides a unique context for learning and experimenting with these technologies in a Do-It-Yourself (DIY) way.

The paper [54] describes the use of the Poppy platform as a tool for scientific researches as well as educationnal and artistic applications.

Several experiments with Poppy in middle and high schools, science museums and Fablabs in France and abroad are already underway and will be discussed in the upcoming **Partnerships and Cooperations** sections.

6.6.3. Expression of emotions with Poppy Humanoid

Participants: Fabien Benureau [correspondant], Matthieu Lapeyre.

Two students in 3rd year of the Cognitive Science major at the University of Bordeaux led a TER project this year using Poppy under the supervision of Fabien Benureau, exploring how the attitude towards robots influences how humans recognise the emotion they try to express. Poppy having no facial expression — or face — yet, the students expressed the five expressions they selected (anger, surprise, joy, sadness, disgust) with body movements alone. They videotaped the sequences of movements (videos are available here http://python.sm.u-bordeaux2.fr/ter/2014/sc/desprez-zerdoumi/?page_id=289) and created an experiment asking volunteers to guess which emotion was displayed. The form also included the Negative Attitude towards Robots Scale (NARS), to investigate the possible correlation between fear of robot and the ability to identify their emotional attitude. The results showed no correlation between the two, although it was admitted that the experiment would have to be improved and ran again before any conclusion could be made.

6.7. Interactive Learning and user adaptation

6.7.1. Interactive learning from unlabeled instructions

Participants: Grizou Jonathan [correspondant], Itturate Inaki, Montesano Luis, Pierre-Yves Oudeyer, Manuel Lopes.

Interactive learning deals with the problem of learning and solving tasks using human instructions. It is common in human-robot interaction, tutoring systems, and in human-computer interfaces such as brain-computer ones. In most cases, learning these tasks is possible because the signals are predefined or an ad-hoc calibration procedure allows to map signals to specific meanings. In this work, we addressed the problem of simultaneously solving a task under human feedback and learning the associated meanings of the feedback signals. This has important practical application since the user can start controlling a device from scratch, without the need of an expert to define the meaning of signals or carrying out a calibration phase. We proposed an algorithm that simultaneously assign meanings to signals while solving a sequential task under the assumption that both, human and machine, share the same a priori on the possible instruction meanings and the possible tasks. This work was published in a conference paper [45] and a journal paper will be submitted in January 2015.

We communicated about this work to the human-robot interaction (HRI) community. A robot equiped with our algortihm would be able to interact with a human without knowing in advance the specific communicative signals used by the human. This work was published in the HRI Pionneer workshop [46].

This work was presented during the thesis defense of Jonathan Grizou entitled: Learning from Unlabeled Interaction Frames, on October 24, 2014. The video, slides, and thesis manuscript can be found at: http://jgrizou.com/projects/thesis-defense/

6.7.2. Calibration-Free BCI Based Control

Participants: Grizou Jonathan [correspondant], Itturate Inaki, Montesano Luis, Pierre-Yves Oudeyer, Manuel Lopes.

We applied previous work on interactive learning from unlabeled instructions [45] to Brain-Machine Interaction problem, leading to a Calibration-Free brain computer interfaces. So far in such brain-computer interfaces (BCI), an explicit calibration phase was required to build a decoder that translates raw electroencephalography signals from the brain of each user into meaningful instructions. Our method removes the calibration phase, and allows a user to control a device to solve a sequential task. We performed experiments where four users use BCI to control an agent on a virtual world to reach a target without any previous calibration process. Our approach is promising for the deployments of BCI applications out of the labs. This work was published in a conference paper [44] and a journal paper will be submitted in January 2015.

This work was presented during the thesis defense of Jonathan Grizou entitled: Learning from Unlabeled Interaction Frames, on October 24, 2014. The video, slides, and thesis manuscript can be found at: http://jgrizou.com/projects/thesis-defense/

6.8. Studying the Co-Construction of Interaction Protocols in Collaborative Tasks with Humans

6.8.1. Experimental Setups for User Study of Alignment in Asymmetric Interactions

Participants: Anna-Lisa Vollmer [correspondant], Jonathan Grizou, Manuel Lopes, Katharina Rohlfing, Pierre-Yves Oudeyer.

In interaction, humans align and effortlessly create common ground in communication, allowing efficient collaboration in widely diverse contexts. Robots are still far away from being able to adapt in such a flexible manner with non-expert humans to complete collaborative tasks. Challenges include the capability to understand unknown feedback or guidance signals, to make sense of what they refer to depending on their timing and context, and to agree on how to organize the interaction into roles and turns.

As a first step in approaching this issue, we investigated the processes used by humans to negotiate a protocol of interaction when they do not already share one. We developed a new experimental setup, where two humans have to collaborate to solve a task. The channels of communication they can use are constrained and force them to invent and agree on a shared interaction protocol in order to solve the task. These constraints allow us to analyze how a communication protocol is progressively established through the interplay and history of individual actions.

We consider a remote construction task, where one user (user A) knows what to build but do not have access to the construction site while its partner (user B) is at the site but do not know what to do. By constraining the communicative channel between the two partners, we study how, and if, they will agree on a similar set of signals to convey information and what type of information they tend to produce. The experimental setup consist of box with button, a video recording system and two screens. User A can send signals to user B by pressing buttons (fig. 14). Signals are displayed on a screen (fig. 14) at user B side. User A is not aware of what is displayed on user B screen, neither user B is aware of the relation between button presses and screen events. The video of user B construction scene is streamed to a screen at user B side. The task consist of bulding arbitrary construction (fig. 14) using colored toy bricks (fig. 14).

The various data recolted during these interaction sequences (fig. 15) allow us to study the Co-Construction of Interaction Protocols. This work was published in a conference paper [66].

6.9. Other

6.9.1. A Framework for Proactive Assistance

Participants: Alexandre Armand, David Filliat [correspondant].

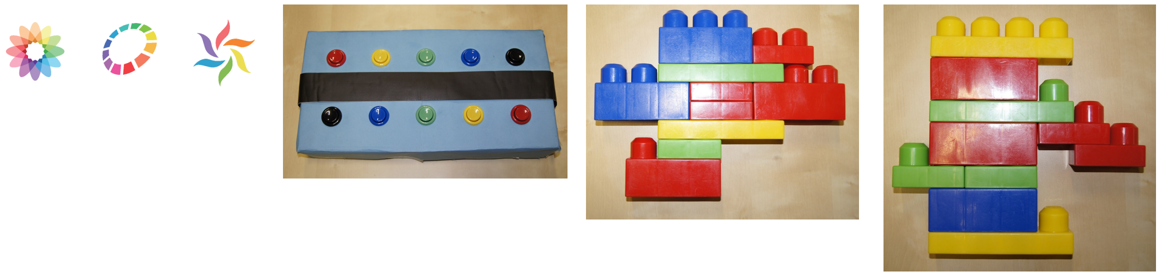


Figure 14. Three examples of sign displayed on the learner screen; The box and the button use as an interface for the teacher to communicate with the learner; Examples of construction presented to the teacher.

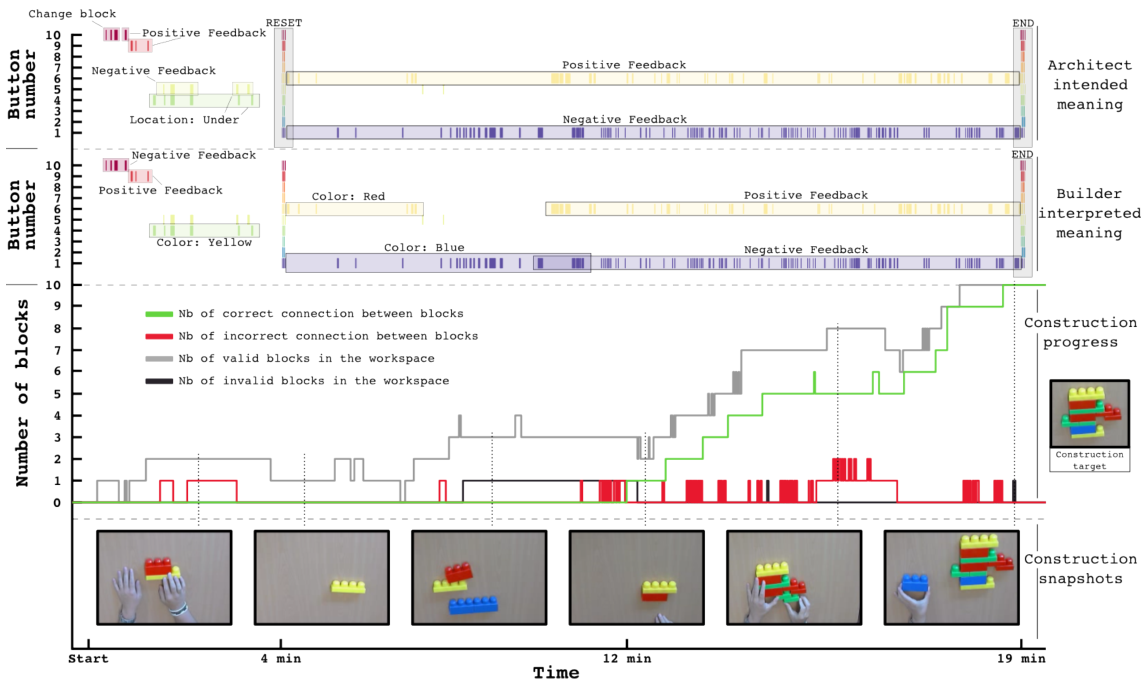


Figure 15. Timeline for one experiment of an architect and a builder collaborating towards building the construction target (right hand side). The top and middle part show the timeline of button presses associated with the intended meaning from the architect (top) and the understood meaning from the builder (middle).

We worked in collaboration with Renault on the problems of adapting driving assitance systems by learning individual drivers behaviours and of integrating more advanced perception in these systems. Advanced Driving Assistance Systems usually provide assistance to drivers only once a high risk situation has been detected. Indeed, it is difficult for an embedded system to understand driving situations, and to predict early enough that it is to become uncomfortable or dangerous. Most of ADAS work assume that interactions between road entities do not exist (or are limited), and that all drivers react in the same manner in similar conditions. We propose a framework that enables to fill these gaps. On one hand, an ontology which is a conceptual description of entities present in driving spaces is used to understand how all the perceived entities interact together with the subject vehicle, and govern its behavior. On the other hand, a dynamic Bayesian Network enables to estimate the driver situation awareness with regard to the perceived objects, based on the ontology inferences, map information, driver actuation and learned driving style. This work was published in a workshop [33] and a conference paper [32].

FLUMINANCE Project-Team

6. New Results

6.1. Highlights of the Year

6.1.1. Stochastic fluid flow dynamics under uncertainty

We have proposed the basis of a formalism allowing to built large scale stochastic representation of fluid flows dynamics [17]. This formalism relies on a location uncertainty principle which separates the flow in terms of a resolved large scale component and a highly oscillating random component. The dynamics is built in a similar way as in the deterministic case through a stochastic representation of the Reynolds transport theorem. This principle paves a new way for the construction of subgrid models from the uncertainties we have on the flow. The associated subgrid tensor provides a clear interaction between small scale data and large scale resolved quantities. This characteristic opens new directions for the devising of methods for the nulmerical simulation of large scale components of the flow. It allows also deriving large-scale models that takes into account explicitly the inherent errors to a particular geophysical dynamics representation.

6.2. Fluid motion estimation

6.2.1. Stochastic uncertainty models for motion estimation

Participants: Etienne Mémin, Manuel Saunier, Abed Malti.

In this study we have proposed a stochastic formulation of the brightness consistency used principally in motion estimation problems. In this formalization the image luminance is modeled as a continuous function transported by a flow known only up to some uncertainties. Stochastic calculus then enables to built conservation principles which take into account the motion uncertainties. These uncertainties defined either from isotropic or anisotropic models can be estimated jointly to the motion estimates. Such a formulation besides providing estimates of the velocity field and of its associated uncertainties allows us to naturally define a linear multiresolution scale-space framework. The corresponding estimator, implemented within a local least squares approach, has shown to improve significantly the results of the corresponding deterministic estimator (Lucas and Kanade estimator). This fast local motion estimator provides results that are of the same order of accuracy than state-of-the-art dense fluid flow motion estimator for particle images. The uncertainties estimated supply a useful piece of information in the context of data assimilation. This ability has been exploited to define multiscale incremental data assimilation filtering schemes. The development of an efficient GPU based version of this estimator recently started through the Inria ADT project FLUMILAB

6.2.2. 3D flows reconstruction from image data

Participants: Ioana Barbu, Kai Berger, Cédric Herzet, Etienne Mémin.

Our work focuses on the design of new tools for the estimation of 3D turbulent flow motion in the experimental setup of Tomo-PIV. This task includes both the study of physically-sound models on the observations and the fluid motion, and the design of low-complexity and accurate estimation algorithms. On the one hand, we investigate state-of-the-art methodologies such as ,“sparse representations" for the characterization of the observation and fluid motion models. On the other hand, we place the estimation problem into a probabilistic Bayesian framework and use state-of- the-art inference tools to effectively exploit the strong time-dependence on the fluid motion.

Last year, we focused on the design of new methodologies to jointly estimate the volume of particles and the velocity field from the received image data. Our approach was based on the minimization (with respect to both the position of the particles and the velocity field) of a cost function penalizing both the discrepancies with respect to a conservation equation and some prior estimates of particle positions.

This year, we revisited the problem of volume reconstruction through the prism of some modern optimization techniques. More specifically, we focussed our attention on the family of proximal and splitting methods and showed that the standard techniques commonly adopted in the TomoPIV literature can be seen as particular cases of such methodologies. Recasting standard methodologies in a more general framework allowed us to propose extensions of the latter: i) we showed that the parcimony characterizing the sought volume can be accounted for without increasing the complexity of the algorithms (e.g., by including simple thresholding operations); ii) we emphasized that the speed of convergence of the standard reconstruction algorithms can be improved by using Nesterov's acceleration schemes; iii) we also proposed a totally novel way of reconstructing the volume by using the so-called "alternating direction of multipliers method" (ADMM) . The journal publications relative to the contributions developped this year are currently in construction.

6.2.3. *Sparse-representation algorithms*

Participant: Cédric Herzet.

The paradigm of sparse representations is a rather new concept which turns out to be central in many domains of signal processing. In particular, in the field of fluid motion estimation, sparse representation appears to be potentially useful at several levels: i) it provides a relevant model for the characterization of the velocity field in some scenarios; ii) it plays a crucial role in the recovery of volumes of particles in the 3D Tomo-PIV problem.

Unfortunately, the standard sparse representation problem is known to be NP hard. Therefore, heuristic procedures have to be devised to access to the solution of this problem. Among the popular methods available in the literature, one can mention orthogonal matching pursuit (OMP), orthogonal least squares (OLS) and the family of procedures based on the minimization of ℓ_p norms. In order to assess and improve the performance of these algorithms, theoretical works have been undertaken in order to understand under which conditions these procedures can succeed in recovering the "true" sparse vector.

Last, we contributed to this research axis by deriving conditions of success for the algorithms mentioned above when some partial information is available about the position of the nonzero coefficients in the sparse vector. This paradigm is of interest in the Tomographic-PIV volume reconstruction problem: one can indeed expect volumes of particles at two successive instants to be pretty similar; any estimate of the position of the particles at one given instant can therefore serve as a prior estimate about their position at the next instant. Another information of interest which can help the algorithms in their reconstruction process is the decay of the amplitude of the nonzero coeffcient in the sparse vector. In a TomoPIV context, this decay corresponds to the fact that not all the particles in fluid diffuse the same quantity of light (notably beacuse of illumination or radius variation). This year, we thus pursue our effort in the understanding of the success of some reconstruction algorithms when the sparse vectors obey some decay. In particular, we showed that the standard coherence-based guarantees for OMP/OLS can be relaxed by an amount which depends on the decay of the nonzero coeffcients.

Another axis of research we have dealt with is the extension of sparse methodologies to the context of nonlinear models. This type of situtation is indeed frequently encountered in fluid mechanics or geophysics where the initial/boundary conditions of a system are known to be sparse in some basis and the collected observations obey a nonlinear dynamical model (e.g., the Navier-Stokes equations). In our work, we showed that many sparse representation algorithms, designed in the linear paradigm, can be nicely extended to the nonlinear setup provided that the gradient of the functional can be evaluated efficiently. In order to do so, we suggested a methodology, well-known in the commmunity of optimal control, but surprinsingly quite uncommon in many fields of signal processing.

Our work have led to the publication of contributions in the IEEE International Conference on Speech, Acoustic and Signal Processing (ICASSP) [23] and international - Traveling Workshop on Interactions between Sparse models and Technology (iTwist) [22],[24]

6.3. Tracking, Data assimilation and model-data coupling

6.3.1. Stochastic filtering technique for the tracking of closed curves

Participant: Etienne Mémin.

We have studied a stochastic filtering technique for the tracking of closed curves along an image sequence. In that goal, we designed a continuous-time stochastic dynamics that allows us to infer inter-frame deformations. The curve is defined by an implicit level-set representation and the stochastic dynamics is expressed properly on the level-set function. It takes the form of a stochastic partial differential equation with a Brownian motion of low dimension. The evolution model we proposed combines local photometric information, deformations induced by the curve displacement and an uncertainty modeling of the dynamics. Specific choices of noise models and drift terms lead to an evolution law based on mean curvature as in classic level set methods, while other choices yield new evolution laws. The approach we propose is implemented through a particle filter, which includes color measurements characterizing the target and the background photometric probability densities respectively. The merit of this parameter free filter is demonstrated on various satellite image sequences depicting the evolution of complex geophysical flows. This work has been recently published in the Journal of Mathematical Imaging and Vision [13]. Let us note the method provides an empirical dynamical model learned recursively from a data flow. Its short time forecasting skills have been used in the context of weather-watch radar images within a fruitful collaboration with MeteoFrance.

6.3.2. Sequential smoothing for fluid motion

Participants: Anne Cuzol, Etienne Mémin.

In parallel to the construction of stochastic filtering techniques for fluid motions, we have proposed a new sequential smoothing method within a Monte-Carlo framework. This smoothing aims at reducing the temporal discontinuities induced by the sequential assimilation of discrete time data into continuous time dynamical models. The time step between observations can indeed be long in environmental applications for instance, and much longer than the time step used to discretize the model equations. While the filtering aims at estimating the state of the system at observations times in an optimal way, the objective of the smoothing is to improve the estimation of the hidden state between observation times. The method is based on a Monte-Carlo approximation of the filtering and smoothing distributions, and relies on a simulation technique of conditional diffusions. The proposed smoother can be applied to general non linear and multidimensional models. It has been applied to a turbulent flow in a high-dimensional context, in order to smooth the filtering results obtained from a particle filter with a proposal density built from an Ensemble Kalman procedure. This conditional simulation framework can also be used for filtering problem with low measurement noise. This has been explored through a collaboration with Jean-Louis Marchand (ENS Bretagne) in the context of vorticity tracking from image data.

6.3.3. Stochastic fluid flow dynamics under uncertainty

Participants: Etienne Mémin, Valentin Resseguier.

In this research axis we aim at devising Eulerian expressions for the description of fluid flow evolution laws under uncertainties. Such an uncertainty is modeled through the introduction of a random term that allows taking into account large-scale approximations or truncation effects performed within the dynamics analytical constitution steps. This includes for instance the modeling of unresolved scales interaction in large eddies simulation (LES) or in Reynolds average numerical simulation (RANS), but also uncertainties attached to non-uniform grid discretization. This model is mainly based on a stochastic version of the Reynolds transport theorem. Within this framework various simple expressions of the drift component can be exhibited for different models of the random field carrying the uncertainties we have on the flow. We aim at using such a formalization within image-based data assimilation framework and to derive appropriate stochastic versions of geophysical flow dynamical modeling. This formalization has been published in the journal Geophysical and Astrophysical Fluid Dynamics [17]. Numerical simulation on divergence free wavelets basis of 3D viscous Taylor-Green vortex and Crow instability have been performed within a collaboration with Souleymane Kadri-Harouna. First promising results have been obtained. Besides, we explore in the context of Valentin Resseguier's PhD the extension of such framework to oceanic models and to satellite image data assimilation. This PhD thesis takes place within a fruitful collaboration with Bertrand Chapron (CERSAT/IFREMER).

6.3.4. Free surface flows reconstruction and tracking

Participants: Dominique Heitz, Etienne Mémin.

We investigated the combined use of a Kinect depth sensor and of a stochastic data assimilation method to recover free-surface flows. More generally, we proposed a particle filter method to reconstruct the complete state of free-surface flows from a sequence of depth images only. The data assimilation scheme introduced accounts for model and observations errors. We evaluated the developed approach on two numerical test cases: a collapse of a water column as a toy-example and a flow in an suddenly expanding flume as a more realistic flow. The robustness of the method to simulated depth data quality and also to initial conditions was considered. We illustrated the interest of using two observations instead of one observation into the correction step. Then, the performance of the Kinect sensor to capture temporal sequences of depth observations was investigated. Finally, the efficiency of the algorithm was qualified for a wave in a real rectangular flat bottom tank. It was shown that for basic initial conditions, the particle filter rapidly and remarkably reconstructed velocity and height of the free surface flow based on noisy measures of the elevation alone. These results have been recently submitted to a special issue of Fluid Dynamics Research.

6.3.5. Variationnal ensemble methods for data assimilation

Participants: Dominique Heitz, Etienne Mémin, Cordelia Robinson, Yin Yang.

In this work, we aim at studying an ensemble based optimal control strategy for data assimilation. Such a formulation nicely combines the ingredients of ensemble Kalman filters and variational assimilation. In the same way as standard variational assimilation, it is formulated as the minimization of an objective function. However, similarly to ensemble filters, it introduces in its objective function an empirical ensemble-based background-error covariance and works in an off-line smoothing mode rather than sequentially like filtering approaches in a sequential filter. These techniques have the great advantage to avoid the constitution of tangent linear and adjoint models, which are necessary for standard incremental variational techniques. As the background error covariance matrix plays a key role in the variational process, our study particularly focuses on the generation of the analysis ensemble state with localization techniques. The proposed method was assessed with a Shallow Water model combined with synthetic data and original incomplete experimental depth sensor observations. Results submitted to Computers & Fluids showed that the modified ensemble technique was better in quality and reduced the computational cost.

6.3.6. Optimal control techniques for the coupling of large scale dynamical systems and image data

Participants: Dominique Heitz, Etienne Mémin, Cordelia Robinson.

This work aims at investigating the use of optimal control techniques for the coupling of Large Eddies Simulation (LES) techniques and 2D image data. The objective is to reconstruct a 3D flow from a set of simultaneous time resolved 2D image sequences visualizing the flow on a set of 2D plans enlightened with laser sheets. This approach will be experimented on shear layer flows and on wake flows generated on the wind tunnel of Irstea Rennes. Within this study we wish also to explore techniques to enrich large-scale dynamical models by the introduction of uncertainty terms or through the definition of subgrid models from the image data. This research theme is related to the issue of turbulence characterization from image sequences. Instead of predefined turbulence models, we aim here at tuning from the data the value of coefficients involved in traditional LES subgrid models or in longer-term goal to learn empirical subgrid models directly from image data. An accurate modeling of this term is essential for Large Eddies Simulation as it models all the non resolved motion scales and their interactions with the large scales.

We have pursued the first investigations on a 4DVar assimilation technique, integrating PIV data and Direct Numerical Simulation (DNS), to reconstruct two-dimensional turbulent flows. The problem we are dealing with consists in recovering a flow obeying Navier-Stokes equations, given some noisy and possibly incomplete PIV measurements of the flow. By modifying the initial and inflow conditions of the system, the proposed method reconstructs the flow on the basis of a DNS model and noisy measurements. The technique has been evaluated in the wake of a circular cylinder. It denoises the measurements and increases the spatiotemporal resolution of PIV time series. These results have been recently published in the Journal of Computational Physics [6]. Along the same line of studies the 3D case is ongoing. The goal consists here to reconstruct a 3D flow from a set of simultaneous time resolved 2D images of planar sections of the 3D volume. This work is mainly conducted within the PhD of Cordelia Robinson. The development of the variational assimilation code has been initiated within a collaboration with A. Gronskis, S. Laizé (lecturer, Imperial College, UK) and Eric Lamballais (institut P' Poitiers). A High Reynolds number simulation of the wake behind a cylinder has been recently performed within this collaboration. The 4DVar assimilation technique based on the numerical code Incompact3D is now implemented. We are currently trying to reconstruct a 3D turbulent flow from dual plane velocity observations. First assessments have been carried out with DNS based synthetic data. Further evaluation will be done with real measurements based on dual stereo PIV experiments.

6.3.7. Ensemble variational data assimilation of large scale fluid flow dynamics with uncertainty

Participants: Etienne Mémin, Yin Yang.

In this work we explore the assimilation of a large scale representation of the flow dynamics with image data of finer resolution. The velocity field at large scales is described as a regular smooth component whereas the complement component is a highly oscillating random velocity field defined on the image grid but living at all the scales. Following this route we have started to assess the performance of a variational assimilation technique with direct image data observation. Preliminary encouraging results have been obtained for simulation under uncertainty of 1D and 2D shallow water models.

6.3.8. Reduced-order models for flows representation from image data

Participants: Cédric Herzet, Etienne Mémin.

One of the possibilities to neglect the influence of some degrees of freedom over the main characteristics of a flow consists in representing it as a sum of K orthonormal spatial basis functions weighted with temporal coefficients. To determine the basis function of this expansion, one of the usual approaches relies on the Karhunen-Loeve decomposition (refered to as proper orthogonal decomposition – POD – in the fluid mechanics domain). In practice, the spatial basis functions, also called modes, are the eigenvectors of an empirical auto-correlation matrix which is built from “snapshots" of the considered physical process.

In this axis of work we focus on the case where one does not have a direct access to snapshots of the considered physical process. Instead, the POD has to be built from the partial and noisy observations of the physical phenomenon of interest. Instances of such scenarios include situations where real instantaneous vector-field snapshots are estimated from a sequence of images. We have been working on several approaches dealing with such a new paradigm. A first approach consists in extending standard penalized motion-estimation algorithms to the case where the sought velocity field is constrained to span a low-dimensional subspace. In particular, we have considered scenarios where the standard optical flow constraint (OFC) is no longer satisfied and one has therefore to resort to a Discrete Finite Difference (DFD) model. The non-linearity of the latter leads to several practical issues that we have addressed this year.

6.4. Analysis and modeling of turbulent flows

6.4.1. Hot-wire anemometry at low velocities

Participant: Dominique Heitz.

A new dynamical calibration technique has been developed for hot-wire probes. The technique permits, in a short time range, the combined calibration of velocity, temperature and direction calibration of single and multiple hot-wire probes. The calibration and measurements uncertainties were modeled, simulated and controlled, in order to reduce their estimated values. Based on a market study the french patent application has been extended this year to a Patent Cooperation Treaty (PCT) application.

6.4.2. Numerical and experimental image and flow database

Participant: Dominique Heitz.

The goal was to design a database for the evaluation of the different techniques developed in the Fluminance group. The first challenge was to enlarge a database mainly based on two-dimensional flows, with three-dimensional turbulent flows. Synthetic image sequences based on homogeneous isotropic turbulence and on circular cylinder wake have been provided. These images have been completed with time resolved Particle Image Velocimetry measurements in wake and mixing layers flows. This database provides different realistic conditions to analyse the performance of the methods: time steps between images, level of noise, Reynolds number, large-scale images. The second challenge was to carried out orthogonal dual plane time resolved stereoscopic PIV measurements in turbulent flows. The diagnostic employed two orthogonal and synchronized stereoscopic PIV measurements to provide the three velocity components in planes perpendicular and parallel to the streamwise flow direction. These temporally resolved planar slices observations will be used in 4DVar assimilation technique, integrating Direct Numerical Simulation (DNS) and Large Eddies Simulation (LES), to reconstruct three-dimensional turbulent flows. This reconstruction will be conducted within the PhD of Cordelia Robinson. The third challenge was to carried out a time resolved tomoPIV experiments in a turbulent wake flow. These temporally resolved volumic observations will be used to assess the algorithms developped in the PhD of Ioana Barbu and in the postdoc of Kai Berger. Then this data will be used in 4DVar assimilation technique to reconstruct three-dimensional turbulent flows. This reconstruction will be conducted within the PhD of Cordelia Robinson.

6.5. Visual servoing approach for fluid flow control

6.5.1. Minimization of the kinetic energy density in the 2D plane Poiseuille flow

Participants: Christophe Collewet, Xuan Quy Dao.

This works concerns the PhD thesis of Xuan-Quy Dao. This year we have focused on a way to ensure a strict decreasing of the kinetic energy density. In that purpose, we have first proposed an approach to increase the controlled degrees of freedom. Indeed, the classical way to model this flow leads to only two degrees of freedom. With so few degrees of freedom it is obviously impossible to reach high desired performances as the strict minimization of the kinetic energy density. This way to proceed leads to a better minimization of the kinetic energy density. We have also proposed on approach based on a local decoupling of the controlled degree of freedom of the system so that an exponential decoupled decrease of each components of the state vector is locally obtained.

6.5.2. Control of systems described by partial differential equations

Participants: Tudor-Bogdan Airimitoaie, Christophe Collewet.

This work concerns principally the post-doctoral research of Tudor-Bogdan Airimiţoaie. It aims at controlling continuously evolving systems described by partial differential equations (PDEs). This is relevant in the context of the Fluminance team because fluid flows are infinite dimensional systems and can be rigorously described only through PDEs. In spite of this, practical approaches of flow control are based on low order numerical implementation relying on space and time discretization of the continuous system. This implies to setup strategies for model reduction that must be then in return properly understood with respect to the convergence of the control law. For finite dimensional implementations, one of the research directions pursued concerns the study on the benefit of increasing the controlled degrees of freedom (see the work of Xuan-Quy Dao). Another research direction, started recently, consists in improving control by using real-time estimation of a finite number of parameters related to the original infinite dimensional system. Indeed, this opens the possibility of improving performances by using more advanced robust linear parametric varying (LPV) control techniques existing in the literature. Two conference papers on these works have been submitted at the "7th AIAA Flow Control Conference".

FOCUS Project-Team

6. New Results

6.1. Highlights of the Year

Valeria Vignudelli has received the AILA (Associazione Italiana di Logica e sue Applicazioni) award for her 2014 master thesis.

6.2. Service-oriented computing

Participants: Maurizio Gabbrielli, Elena Giachino, Saverio Giallorenzo, Claudio Guidi, Mario Bravetti, Ivan Lanese, Michael Lienhardt, Jacopo Mauro, Fabrizio Montesi, Gianluigi Zavattaro.

6.2.1. Orchestrations

Orchestration models and languages in the context of Service-Oriented Architectures (SOA) are used to describe the composition of services focusing on their interactions. Coloured Petri nets (CPN) offer a formal yet easy tool for modelling interactions in SOAs, however mapping abstract SOAs into executable ones requires a non-trivial and time-costly analysis. In [34], we propose a methodology that maps CPN-modelled SOAs into Jolie SOAs (our target language), exploiting a collection of recurring control-flow patterns, called Workflow Patterns, as composable blocks of the translation. We validate our approach with a realistic use case. In addition, we pragmatically assess the expressiveness of Jolie with respect to the considered WPs.

6.2.2. Choreographies

Choreographies are high-level descriptions of distributed interacting systems featuring as basic unit a communication between two participants. A main feature of choreographies is that they ensure deadlock-freedom by construction. From a choreography one can automatically derive a description of the behaviour of each participant using a notion of projection. Choreographies can be used both at the level of types (multiparty session types) or as a programming language. In [18] we surveyed the work on choreographies and behavioural contracts in multiparty interactions. In [28] we explored the notion of deadlock freedom (the system never gets stuck), and the related notions of lock freedom (each action is eventually executed under a fair scheduling) and progress (each session never gets stuck). Previous work studied how to define progress in an open setting by introducing the notion of catalysers, execution contexts generated from the type of a process. We refined the notion of catalysers leading to a novel characterization of progress in terms of the standard notion of lock-freedom. We applied our results both to binary session types and in an untyped session-based setting. We combined our results with existing techniques for lock-freedom, obtaining a new methodology for proving progress. Our methodology captures new processes w.r.t. previous progress analysis based on session types. The two following works consider the extension of choreographies, which traditionally have a static structure, to deal with adaptation, i.e., dynamic changes of the structure of choreographies. A preliminary analysis of adaptable choreographies at the level of types is presented in [27]. This work considers both updates from inside the system (self-adaptation), and external updates. Adaptable choreographies as a programming language are considered in [33], where we presented AIOCJ, a framework for programming distributed adaptive applications. AIOCJ allows the programmer to specify which parts of the application can be adapted. Adaptation takes place at run-time by means of rules, which can change during the execution to tackle possibly unforeseen adaptation needs. AIOCJ relies on a solid theory that ensures applications to be deadlock free by construction also after adaptation.

6.3. Models for reliability

Participants: Mario Bravetti, Elena Giachino, Ivan Lanese, Michael Lienhardt, Gianluigi Zavattaro.

6.3.1. Reversibility

We have continued the study of causal-consistent reversibility started in the past years. In [17] we presented an overview of causal-consistent reversibility, summarizing the main approaches in the literature, and the related results and applications. An interesting application is debugging. Reversible debugging provides developers with a way to execute their applications both forward and backward, seeking the cause of a misbehaviour. In a concurrent setting, reversing actions in the exact reverse order they have been executed may lead to undo many actions that were not related to the bug under analysis. On the other hand, undoing actions in some order that violates causal dependencies may lead to states that could not be reached in a forward execution. In [36] we proposed a new approach, where each action can be reversed if all its consequences have already been reversed. The main feature of the approach is that it allows the programmer to easily individuate and undo exactly the actions that caused a given misbehaviour till the corresponding bug is reached. We discussed the appropriate primitives for causal-consistent reversible debugging and presented their prototype implementation in the CaReDeb tool.

6.3.2. Fault models

We have continued the study of primitives for fault handling in a concurrent setting. In [19] we critically discussed the different choices that have to be made when defining a fault model for a concurrent object-oriented programming language. We consider in particular the ABS language, and analyse the interplay between the fault model and the main features of ABS, namely the cooperative concurrency model, based on asynchronous method invocations whose return results via futures, and its emphasis on static analysis based on invariants.

6.4. Cloud Computing

Participants: Roberto Amadini, Maurizio Gabbrielli, Elena Giachino, Saverio Giallorenzo, Claudio Guidi, Cosimo Laneve, Michael Lienhardt, Tudor Alexandru Lascu, Jacopo Mauro, Gianluigi Zavattaro.

6.4.1. Cloud application deployment

Configuration and management of applications in the cloud is a complex task that requires novel methodologies and tools. In [16] we have performed a foundational study of the complexity boundaries for the automatic deployment problem, showing that in the general case this problem is undecidable, it is decidable but non-primitive recursive if capacity constraints are not taken into account, while it turns out to be polynomial time if also conflicts between software components are not considered. Starting from these foundational observations, we have investigated the exploitability in this specific context of state-of-the-art constraint optimization techniques, a well established approach for the modeling and solution of complex optimization problems. In particular, in [23], [24] we have studied how the "portfolio technique" approach can be applied to optimization problems, combining and exploiting the performances of existing solvers to get a global, more robust and fast solver. Encouraged by these results, we have developed SUNNY-CP [13], [22]: a porftolio constraint solver for constraint satisfaction and optimization problems. SUNNY-CP has proven to have remarkable performances, ranking 4th in the annual MiniZinc challenge (i.e., the international competition to evaluate the performances of constraint solvers) and receiving a ‘honorable’ mention by the challenge organizers.

6.4.2. Cloud resource management

The management of cloud resources from client programs requires the definition of Application Programming Interfaces (APIs) that expose specific functionalities to external invokers. Programs can be built that compose existing APIs in order to obtain new functionalities. However API composition easily becomes a frustrating and time-costly task that hinders API reuse. The issue derives from technology-dependent features of API composition such as the need of extensive documentation, protocol integration, security issues, etc.. In [39] we introduce the perspective of the API-as-a-Service (APIaaS) layer as tool to ease the development and deployment of applications based on API composition, abstracting communication protocols and message formats. We elicit the desirable features of such a layer and provide a proof-of-concept prototype implemented using a service-oriented language.

Another critical aspect in this context deals with the problem of dynamic reallocation of resources. In [38] we study a type-based technique for modeling and analysis of systems in which concurrent object-oriented programs dynamically create and move resources. The type of a program is behavioural, namely it expresses the resource deployments over periods of (logical) time. Our technique admits the inference of types and may underlie the optimisation of the costs and consumption of resources.

6.5. Resource Control

Participants: Michele Alberti, Alberto Cappai, Ugo Dal Lago, Simone Martini, Giulio Pellitta, Davide Sangiorgi, Marco Solieri, Valeria Vignudelli.

6.5.1. Probabilistic higher-order calculi

The first results of our efforts on probabilistic higher-order systems and languages have started to appear in 2014. In particular, we have focused our attention on the impact of probability to the classical notion of context equivalence for the lambda-calculus, showing that applicative bisimilarity continues to be a congruence [31], and that it even coincides with context equivalence when evaluation is done in the call-by-value order [29]. The expressive power of higher-order concurrent contexts has been compared to the expressive power of lambda-calculi contexts and put in relation with other equivalences when the observed process is either an ordinary Labelled Transition Systems (LTS) or a reactive probabilistic transition system [25]. The obtained spectrum of equivalences for reactive probabilistic processes has been shown to be finer than the one for classic LTSs. We have also analysed the expressive power of different first-order testing equivalences (with nondeterministic tests, probabilistic tests, and both nondeterministic and probabilistic tests) in the spectrum for reactive probabilistic processes [26].

6.5.2. Resource consumption

The main result about resource consumption has been about an open problem on the λ-calculus: we proved that the number of leftmost-outermost steps to normal form is indeed an invariant cost model in the sense of Slot and van Emde Boas' weak invariance thesis [21]. We also introduced a new recursion theoretic framework for probabilistic computation in which one is able to capture probabilistic polynomial time through Leivant's Tiering [32].

6.5.3. Geometry of interaction

Novel results have been obtained for Geometry of Interaction (GoI), itself a semantics framework for linear logic introduced by Jean-Yves Girard thirty years ago. In particular, we have shown how the most concrete presentations of GoI, namely so-called token machines, can go *parallel*, thus exploiting the potential parallelism in functional programs (through the Curry-Howard Correspondence). This has been made concrete by studying extensions of multiplicative linear logic in which synchronization becomes an operator where tokens can indeed synchronize [30]. This has been later shown to be necessary to model quantum computation [44]. A simple, minimalistic GoI model of the resource λ-calculus has also been introduced [43].

6.6. Verification techniques for extensional properties

Participants: Daniel Hirschkoff, Elena Giachino, Michael Lienhardt, Cosimo Laneve, Jean-Marie Madiot, Davide Sangiorgi.

Extensional refers to properties that have to do with behavioural descriptions of a system (i.e., how a system looks like from the outside). Examples of such properties include classical functional correctness and deadlock freedom. Related to techniques for extensional properties are the issues of decidability (the problem of establishing whether certain properties are computationally feasable).

6.6.1. Coinductive techniques

Coinductive techniques, notably those based on bisimulation, are widely used in concurrency theory to reason about systems of processes. The bisimulation proof method can be enhanced by employing 'bisimulations up-to' techniques. A comprehensive theory of such enhancements has been developed for first-order (i.e., CCS-like) LTSs and bisimilarity, based on the notion of compatible function for fixed-point theory. We have transported this theory onto languages whose bisimilarity and LTS go beyond those of first-order models [40]. The approach consists in exhibiting fully abstract translations of the more sophisticated LTSs and bisimilarities onto the first-order ones. This allows us to reuse directly the large corpus of up-to techniques that are available on first-order LTSs. We have investigated the method on the π-calculus, the Higher-Order π-calculus, and a (call-by-value) λ-calculus with references.

In [20], mostly a tutorial paper, a few forms of bisimulation and of coinductive techniques that have been proposed for higher-order languages are discussed, beginning with the pure lambda-calculus and then moving to extensions of it, notably those with non-determinism and probabilities.

6.6.2. Deadlock detection

Deadlock detection in concurrent programs that create networks with an arbitrary number of nodes is extremely complex and solutions either give imprecise answers or do not scale. To enable the analysis of such programs, we have studied an algorithm for detecting deadlocks [37], [35], in a basic model featuring recursion and fresh name generation, called Lam. We then have designed a type system that associates Lams to processes. As a byproduct of these two techniques, we have an algorithm that is more powerful than previous ones and that can be easily integrated into the current release of TyPiCal, a type-based analyser for π-calculus.

6.6.3. Expressiveness and decidability in actor-like systems

Refining work in previous years, we have studied [15] the expressive power of an actor-like language, featuring concurrent objects and asynchronous message-passing. We have identified the presence/absence of fields as a crucial feature: the dynamic creation of names in combination with fields gives rise to Turing completeness. On the other hand, restricting to stateless actors gives rise to systems for which properties such as termination are decidable. This decidability result still holds for actors with states when the number of actors is bounded and the state is read-only.

FUN Project-Team

5. New Results

5.1. Highlights of the Year

- Opening of the 256 M3 sensor nodes of the Lille's FIT IoT Lab platform.
- We have designed a novel single-based localization method, UNS, for accurate localization of mobile devices that only needs a small aperture array unlike all previous works. UNS is currently under patenting.
- We have provided a set of recognized contributions in the area of Smart Cities, re-thinking their architecture and break vertical silos between every network and application.

5.2. Routing in FUN

Participants: Valeria Loscri, Nathalie Mitton, Riccardo Petrolo.

According to a wide range of studies, IT should become a key facilitator in establishing primary education, reducing mortality and supporting commercial initiatives in Least Developed Countries (LDCs). The main barrier to the development of IT services in these regions is not only the lack of communication facilities, but also the lack of consistent information systems, security procedures, economic and legal support, as well as political commitment. In [3], [10], we propose the vision of an infrastructure-less data platform well suited for the development of innovative IT services in LDCs. We propose a participatory approach, where each individual implements a small subset of a complete information system thanks to highly secure, portable and low-cost personal devices as well as opportunistic networking, without the need of any form of infrastructure. We review the technical challenges that are specific to this approach. Relying on such an infrastructure, wireless routing must be opportunistic and take advantages of the availability of every infrastructure point when in range. Two different approaches depending on the available devices are presented in [20] and [2]. When partial positions of nodes are available, the system can take advantage of such knowledge to enhance the routing performance. This is what has been investigated in[12] where coordinates are used in an opportunistic fashion when available.

5.3. Self-organization

Participants: Natale Guzzo, Valeria Loscri, Nathalie Mitton.

Self-organization encompasses several mechanisms. This year, the FUN research group has contributed to specific aspects; topology importance and clustering.

5.3.1. Impact of the topology

Wireless Sensor Networks (WSN) are composed of constrained devices and deployed in unattended and hostile environments. Most papers presenting solutions for WSN evaluate their work over random topologies to highlight some of their "good" performances. They rarely study these behaviors over more than one topology. Yet, the topology used can greatly impact the routing performances. [13] presents a study of the impact of the network topology on algorithm performance in WSNs and illustrate it with the geographic routing. Geographic routing relies on node coordinates to route data packets from source to destination. We measure the impact of different network topologies from realistic ones to regular and very popular ones through extensive simulation and experimentation campaigns. We show that different topologies can lead to a difference of up to 25% on delivery ratio and average route length and more than 100% on energy costs.

5.3.2. Clustering

Clustering in wireless sensor networks is an efficient way to structure and organize the network. It aims to identify a subset of nodes within the network and bind it a leader (i.e. cluster-head). This latter becomes in charge of specific additional tasks like gathering data from all nodes in its cluster and sending them by using a longer range communication to a sink or a Base Station (BS) which may be far away from the monitoring area. Many algorithms proposed in the literature compute the routing process by clustering the network and by designing new election mechanisms in which the cluster-heads are chosen taking account of the remaining energy, the communication cost and the density of nodes. However, they do not consider the connectivity to the BS, and assume that all the nodes or only few prefixed nodes are able to directly communicate with it. We believe that this assumption is not suitable for many applications of WSN and to tackle this problem we propose CESAR [14], a multi-hop and energy-efficient routing protocol for large-scale WSN which includes a new cluster-head selection mechanism aware of the battery level and the connectivity to the BS. Furthermore, our solution employs an innovative hybrid approach to combine both clustering and on-demand techniques in order to provide an adaptive behavior for different dynamic topologies. Simulation results show that our solution outperforms in terms of energy consumption and data delivery other known routing algorithms in the literature. Note that CESAR is currently the object of two pending patents.

5.4. Controlled mobility based services

Participants: Emilio Compagnone, Valeria Loscri, Karen Miranda, Nathalie Mitton, Tahiry Razafindralambo, Dimitrios Zormpas, Jean Razafimandimby Anjalalaina.

Sensors have more and more functionality in terms of capture techniques, communication capabilities, processing capabilities and energy harvesting. Another interesting feature available on sensors is mobility. The FUN research group tries to exploit the controlled mobility of sensors to solve some known issues in wireless sensors networks regarding deployment or routing but also raises some new challenges regarding coverage optimization and energy harvesting.

5.4.1. Coverage

Wireless sensors are used to gather information from a field of interest. In order to capture all the events in this field, the sensors must be properly placed. When the sensors have motion capabilities such as robots, the deployment can be optimized. The use of controlled mobility raises some new challenges and opportunities in the field of wireless sensor networks. Milan Erdelj and Karen Miranda in [33] presents the advances in context. They provide a detailed literature review regarding the techniques behind controlled mobility in order to deploy or redeploy sensors. When the wireless sensors are mobile, it is possible to optimize the capture of information regarding their time and space evolution. This allows the sensors to focus on different zones of interest depending on the evolution of the observed events. Valeria Loscri, Enrico Natalizio and Nathalie Mitton present a performance evaluation of different algorithms for zone of interest coverage in [18]. Their work particularly focuses on providing a set of distributed version of a combined particle swarm optimization and virtual forces algorithm. The proposed algorithms and their evaluation show an high reactivity to changing events and targets. Energy is an important constraint in wireless sensor networks and message exchange is a functionality that drains huge amount of energy. Dimitrios Zorbas and Christos Douligeris in [30] present a low-overhead localized algorithm for the target coverage problem in wireless sensor networks. To tackle this problem they propose two variations of a localized algorithm with low communication complexity in term of message exchange. The results show a great improvement in terms of communication cost while achieving an adequate network lifetime.

5.4.2. Connectivity and performance

Information gathered by sensors are to be processed in a remote location. The transportation from the point where the raw data is generated (the sensor) and the data processing unit (sink or other infrastructure) relies on routing techniques. Routing is a fundamental functionality of a wireless sensors network. Nicolas Gouvy, Nathalie Mitton and David Simplot-Ryl in their book chapter [34] provide a review of the routing techniques described in the literature. They highlight the challenges, main issues and future work direction in this domain and provide some important assumption and characteristics that should be kept in mind when designing routing protocols for wireless sensor networks. When route between a source and the destination of data does not exist or cannot be established, using a mobile router is a possible solution. Christos Katsikiotis, Dimitrios Zorbas and Periklis Chatzimisios in [15] propose an algorithm that restores connectivity by the use of mobile wireless router after a routing failure. They provide a fast mechanism to heal the network and restore connectivity between the network partitions. In their solution, a mobile wireless router finds the end points that should be re-connect and place itself in the correct position to restore the connectivity. Their solution shows a fast restoration process based on the implementation done on a real robotic platform.

5.4.3. Energy suppliance

Energy is an important constraint in static wireless sensor networks and even more important when sensors are mobile. However, when sensors have motion capabilities, they can use this ability to move toward a recharging point in order to increase the network operation. Dimitrios Zorbas and Tahiry Razafindralambo in [31] use the motion capability of sensors to provide an algorithm that allow the sensor to go to a recharging point while minimize the impact of their movement on the network operation such as portioning or data gathering. They provide theoretical bounds on the realisation of such operation and evaluate the average behaviour of their algorithm based on extensive simulations. Both results show a big improvement in terms of network lifetime extension compared to the case where no replacement is performed and to the case where rerouting is considered.

5.4.4. Video-based applications

Video Surveillance and Target Detection represent key components for many organizations in terms of safety and security protocols. The value of Video Surveillance has become more sophisticated and very accurate, by leveraging specific sensors able to detect motion, heat, etc. In [17], Valeria Loscri, Michele Magno and Rosario Surace show how the nodes of a sensor network can learn which is their best position based on a certain number of WebCams that need to be "woken-up" when a suspicious event is detected. The main purpose is to reduce power consumption, especially in the case of Video Surveillance, when the most of the time the power is wasted by doing nothing. On the other hand, Target Detection, namely determining whether or not a target object exists in a video frame, has grown significantly with the recent advances in embedded computing and sensors which have opened the possibility to realize smaller and low-cost autonomous systems. In [16], Valeria Loscri, Nathalie Mitton and Emilio Compagnone show the feasibility of low-cost embedded system for detection of objects based either on the shape or on the color.

5.5. Security

Participants: Valeria Loscri, Nathalie Mitton.

Security has been always a critical issue both for the users and providers of wireless communication systems. The definition of novel paradigms and innovative communication systems, such as the Internet of Things (IoT) and the nanocommunication systems, exacerbated the criticality of security and privacy factors. These latter aspects are faced in [23] and [5]. In [23], Riahi et al. face with the security issues related to the IoT paradigm, by taking into consideration that this paradigm enable daily objects to become active participants of everyday activities. They envisage the main challenges and propose solutions to address them. In [5], Valeria Loscri et al. analyze the innovative aspects that characterize the molecular communication paradigm, by proposing innovative and revolutionary methods that take into consideration the very limited available resources (i.e. we work at molecular level and then we cannot leverage on high processing and computing capabilities) and the very high criticality of the potential applications of similar systems (e.g. in-vivo applications).

5.6. RFID

Participants: Ibrahim Amadou, Nathalie Mitton.

Due to the dedicated short range communication feature of passive radio frequency identification (RFID) and the closest proximity operation of both tags and readers in a large-scale dynamic RFID system, when nearby readers simultaneously try to communicate with tags located within their interrogation range, serious interference problems may occur. Such interferences may cause signal collisions that lead to the reading throughput barrier and degrade the system performance. Although many efforts have been done to maximize the throughput by proposing protocols such as NFRA or more recently GDRA, which is compliant with the EPCglobal and ETSI EN 302 208 standards. However, the above protocols are based on unrealistic assumptions or require additional components with more control packet and perform worse in terms of collisions and latency, etc. In [9], we explore the use of some well-known Carrier Sense Multiple Access (CSMA) backoff algorithms to improve the existing CSMA-based reader-to-reader anti-collision protocol in dense RFID networks. Moreover, the proposals are compliant with the existing standards. We conduct extensive simulations and compare their performance with the well-known state-of-the-art protocols to show their performance under various criteria. We find that the proposals improvement are highly suitable for maximizing the throughput, efficiency and for minimizing both the collisions and coverage latency in dense RFID Systems.

5.7. VANET

Participant: Nathalie Mitton.

VANET (Vehicular Networks) is an arising kind of network which features specific functionalities and requirements especially in terms of delay.

[26] analyzes the information delivery delay for roadside unit deployment in an intermittently connected vehicular network. A mathematical model is developed to describe the relationship between the average information delivery delay and the distance between two neighbor RSUs (Road Side Unit) deployed along a road. The derived model considers a straight highway scenario where two RSUs are deployed at a distance without any direct connection and vehicles are sparsely distributed on the road with road condition information randomly generated between the two neighbor RSUs. Moreover, the model takes into account the vehicle speed, the vehicle density, the likelihood of an incident, and the distance between two RSUs. The effectiveness of the derived mathematical model is verified through simulation results. Given the delay requirement of some time-critical applications, this model can be used to estimate the maximum distance allowed between two neighbor RSUs, which can provide a reference basis for the deployment of RSUs in such scenarios.

Abstract–Broadcasting is an effective routing paradigm for data dissemination in vehicular ad hoc networks (VANETs). One concern that arises with broadcasting is the broadcast storm problem, which would cause node contentions and data collisions, and thus degrade the transmission efficiency of a network. [27] proposes a Dynamic trAnsmission delaY based broadcast (DAYcast) protocol for a VANET. To alleviate the effect of the broadcast storm and improve the transmission efficiency of the network, DAYcast only allows the effective neighbors of a source vehicle to broadcast a received data packet and the selection of the effective neighbors are based on the position information on the one-hop neighbors of the source vehicle. Meanwhile, it allows each effective neighbor to wait a certain transmission delay before it broadcasts a received packet. The transmission delay of an effective neighbor depends on the distance between the neighbor and the source vehicle, and the number of effective neighbors of the source vehicle. Simulation results show that DAYcast can effectively improve the network performance in terms of network reachability and the successful delivery ratio as compared with existing weighted p-persistence broadcasting (WPB) and slotted 1-persistence broadcasting (SPB).

5.8. Smart cities architecture

Participants: Valeria Loscri, Nathalie Mitton, Riccardo Petrolo, Nicola Zema.

Smart City represents one of the most promising and prominent Internet of Things (IoT) applications. In the last few years, indeed, smart city concept has played an important role in academic and industry fields, with the development and deployment of various middleware platforms. However, this expansion has followed distinct approaches creating, therefore, a fragmented scenario, in which different IoT ecosystems are not able to communicate between them. To fill this gap, there is a need to re-visit the smart city IoT semantic and offer a global common approach. In order to allow cities to share data across systems and coordinate processes across domains, it is essential to break these silos. A way to achieve the purpose is sensor virtualization, discovery and data restitution. This last year, the FUN team has lead several investigations in this direction.

We have looked at the heterogeneity of devices and network technologies under a different perspective by not perceiving it as a limitation but as a potential to increase the connectivity in a smart city [1]. We propose a new generation of network nodes, called stem nodes, based on the innovative idea of 'stemness', which pushes forward the well-known self-configuration and self-management concepts towards the idea of node mutation and evolution. We also deployed prototypes that demonstrate the stem-node architecture and basic operations in different hardware platforms of common communication devices (an Alix-based router, a laptop and a smartphone)

In [7], we illustrate semantic interoperability solutions for IoT systems. Based on these solutions, we describe how the FP7 VITAL project aims to bridge numerous silo IoT deployments in smart cities through repurposing and reusing sensors and data streams across multiple applications without carelessly compromising citizens' security and privacy. This approach holds the promise of increasing the Return-On-Investment (ROI), which is associated with the usually costly smart city infrastructures, through expanding the number and scope of potential applications.

To this purpose, [21] browses the semantic annotation of the sensors in the cloud, and innovative services can be implemented and considered by bridging Clouds and Internet of Things. Things-like semantic will be considered to perform the aggregation of heterogeneous resources by defining the Clouds of Things paradigm. We survey the smart city vision, providing information on the main requirements and highlighting the benefits of integrating different IoT ecosystems within the cloud under this new CoT vision. This paper also discusses relevant challenges in this research area.

Going further, we also presented [22] a first implementation of this federation: a federation of FIT IoT-LAB within OpenIoT. OpenIoT is a middleware that enables the collection of data streams from multiple heterogeneous geographically dispersed data sources, as well as their semantic unification and streaming with a cloud infrastructure. Future Internet of Things IoT-LAB (FIT IoT-LAB) provides a very large scale infrastructure facility suitable for testing small wireless sensor devices and heterogeneous communicating objects. The integration proposed represents a way to reduce the gap existing in the IoT fragmentation, and, moreover, allows users to develop smart city applications by interacting directly with sensors at different layers. We illustrate it trough a basic temperature monitoring application to show its efficiency.

So, once all city network and infrastructure are set at the same level thanks to the above mentioned approaches, they can go further and offer additional services. An example of them is navigation[11] as also described in "Localization" section. Another example is to make use of the urban bikes [19]. Indeed, besides the growing enthusiast provoked by bicycles in smart and green cities and the benefit for health they bring, there still exists some reluctance in using bikes because of safety, road state, weather, etc. To counter-balance these feelings, there is a need to better understand bicycle users habits, path, road utilization rate in order to improve the bicycle path quality. In this perspective,in this paper, we propose to deploy a set of mobile sensors on bicycles to gather this different data and to exploit them to make the bike easier and make people want to ride bicycles more often. Such a network will also be useful for several entities like city authorities for road maintenance and deployment, doctors and environment authorities, etc. Based on such a framework,we propose a first basis model that helps to dimension the network infrastructure and the kind of data to be real time gathered from bikes. More specifically, we present a theoretical model that computes the quantity of data a bike will be able to send along a travel and the quantity of data a base station should be able to absorb.We have based our study on real data to provide first numerical results and be able to draw some preliminary conclusions and open new research directions.

5.9. Localization

Participants: Ibrahim Amadou, Roudy Dagher, Nathalie Mitton, Roberto Quilez, Nicola Zema.

Navigate in or based on a wireless sensor network present many advantages but it is still an open issue. We have focused on two particular cases in which navigation or WSN-based localization is needed [32]. The former aspect considers that sensors need to be visited on-demand by a mobile sink to offload data. This mobile sink thus needs to locate the data source. The second aspect feature a mobile entity that is needed to be localized.

In a event-based WSN, where is necessary a prompt response in terms of data processing and offloading, a set of mobile flying sinks could be a good option for the role of autonomous data collectors. For those reasons in [28], we propose a distributed algorithm to independently and autonomously drive a mobile sink through the nodes of a WSN and we show its preferability over more classical routing approaches especially in the presence of a localized generation of large amount of information. Our result shows that, in the case of fairly complete coverage of the area where the nodes lie, it is possible to promptly notify a mobile sink about the presence of data to offload, drive it to the interested area and achieve interesting performances. [29] enhanced the previous approach by relaxing some GPS-use assumptions. We show that, under fairly common circumstances, it is possible to set the trajectory of the mobile sink and fulfill the offloading requests without the needs of additional equipment installed on nodes. We show how our system is preferable over more classical routing solutions especially in the presence of localized generation of large amounts of information.

[11] proposes Ubiquitous Navigation System (UNS), a WSN-based navigation system, which takes benefit from a WSN mesh deployment to provide a local navigation service. The positioning part of the system uses Angle of Arrival (AoA) measurements to estimate the vehicle position on the map. Based on a realistic network scenario, extracted from a city map using Google Maps, we study the performance of Triangulation using AoA in a smart urban environment that exhibits topology related constraints. Simulations results show that such constraints lead to particular spatial distribution of the anchor nodes that affects both positioning accuracy and beacon packets reception rate. We also propose and evaluate the use of the network communication range as a technique to mitigate the effect of geometric dilution of precision (GDOP). The simulation results show that this technique successfully detected GDOP-affected positions and thus significantly enhanced the positioning accuracy. One of the biggest strengths of UNS is that it relies on a single anchor unlike literature approaches. The different underlying studies are detailed in [38] in which we study the ambiguity of source localization using signal processing of large aperture antenna arrays under spherical wave propagation. This novel localization approach has been recently proposed, providing an estimate of the source position by means of two methods: geometrical and analytical. The former finds the source position as the estimate of circular loci, the latter as a solution of a linear system of equations. Although this method is proved to work for a general array geometry, we show that it suffers from ambiguities for a particular class of array geometries. Namely, in 2D, we prove that when the array geometry is linear or circular, there exist two possible solutions where only one corresponds to the actual position of the source. We also prove a relation of symmetry between the solutions with respect to the array geometry. This relation is very useful to assist the disambiguation process for discounting one of the estimates. By extension to 3D, planar (resp. spherical) arrays exhibit the same behavior i.e they provide two symmetrical estimates of the source position when the latter is not on the array plane (resp. sphere).
Note that UNS is currently a pending patent.

GALAAD2 Team

6. New Results

6.1. Algebraic representations for geometric modeling

6.1.1. A comparison of different notions of ranks of symmetric tensors

Participants: Alessandra Bernardi, Jérome Brachat, Bernard Mourrain.

In [2], we introduce various notions of rank for a symmetric tensor, namely: rank, border rank, catalecticant rank, generalized rank, scheme length, border scheme length, extension rank and smoothable rank. We analyze the stratification induced by these ranks. The mutual relations between these stratifications, allow us to describe the hierarchy among all the ranks. We show that strict inequalities are possible between rank, border rank, extension rank and catalecticant rank. Moreover we show that scheme length, generalized rank and extension rank coincide.

6.1.2. Dimensions and bases of hierarchical tensor-product splines

Participant: Bernard Mourrain.

In [1], we prove that the dimension of trivariate tensor-product spline space of tri-degree (m,m,m) with maximal order of smoothness over a three-dimensional domain coincides with the number of tensor-product B-spline basis functions acting effectively on the domain considered. A domain is required to belong to a certain class. This enables us to show that, for a certain assumption about the configuration of a hierarchical mesh, hierarchical B-splines span the spline space. This paper presents an extension to three-dimensional hierarchical meshes of results proposed recently by Giannelli and Jüttler for two-dimensional hierarchical meshes.

Joint work with Dmitry Berdinsky, Taiwan Kim, Oh Min-Jae, Sutipong Kiatpanichgij (Department of Naval Architecture and Ocean Engineering, Seoul, South Korea), Cesare Bracco (Dipartimento di Matematica "Giuseppe Peano", Torino, Italy), Durkbin Cho (Department of Mathematics, Dongguk, South Korea).

6.1.3. Bounds on the dimension of trivariate spline spaces: A homological approach

Participant: Bernard Mourrain.

In [8], we consider the vector space of globally differentiable piecewise polynomial functions defined on a three-dimensional polyhedral domain partitioned into tetrahedra. We prove new lower and upper bounds on the dimension of this space by applying homological techniques. We give an insight of different ways of approaching this problem by exploring its connections with the Hilbert series of ideals generated by powers of linear forms, fat points, the so-called Fröberg–Iarrobino conjecture, and the weak Lefschetz property.

Joint work with Nelly Villamizar (RICAM - Johann Radon Institute for Computational and Applied Mathematics, Linz, Austria)

6.1.4. High-quality construction of analysis-suitable trivariate NURBS solids by reparameterization methods

Participants: André Galligo, Bernard Mourrain.

High-quality volumetric parameterization of computational domain plays an important role in three-dimensional isogeometric analysis. Reparameterization techniques can improve the distribution of isoparametric curves/surfaces without changing the geometry. In [10], using the reparameterization method, we investigate the high-quality construction of analysis-suitable NURBS volumetric parameterization. Firstly, we introduce the concept of volumetric reparameterization, and propose an optimal Möbius transformation to improve the quality of the isoparametric structure based on a new uniformity metric. Secondly, from given boundary NURBS surfaces, we present a two-stage scheme to construct the analysis-suitable volumetric parameterization: in the first step, uniformity-improved reparameterization is performed on the boundary surfaces to achieve high-quality isoparametric structure without changing the shape; in the second step, from a new variational harmonic metric and the reparameterized boundary surfaces, we construct the optimal inner control points and weights to achieve an analysis-suitable NURBS solid. Several examples with complicated geometry are presented to illustrate the effectiveness of proposed methods.

Joint work with Gang Xu (College of computer - Hangzhou Dianzi University, China), Timon Rabczuk (Bauhaus-Universität Weimar, Germany).

6.1.5. Spline Spaces over Quadrangle Meshes with Complex Topologies

Participants: André Galligo, Bernard Mourrain, Meng Wu.

Motivated by Magneto Hydrodynamic (MHD) simulation with isoparametric elements method, we pursue our work on new types of spline functions defined over a quadrangular mesh, that can follow isobaric curves with node singularities. The practicability of these splines is analyzed for different geometries related to MHD simulation.

This work is done in collaboration with Boniface Nkonga (Inria, EPI CASTOR and University of Nice).

6.1.6. Parametric modeling for ship hull deformation

Participant: Elisa Berrini.

The objective of the work is to develop a parametric modeler tool, allowing consistent ship hull deformations with respect to classic naval architecture design constraints. This work will be applied in automatic shape optimization process. Two scientific problematics are addressed : 1) The parametrization of the hull: the numerical representation of the shape from a defined set of parameters; 2) The deformations of curves and surfaces: getting a new shape by modifying chosen parameters from the parameterization set. The consistency with naval architecture constraints is essential.

To produce realistic models, we want to use methods similar to naval architects' ones. The approach under development is based on the extraction and deformation of skeletons curves.

6.2. Algebraic algorithms for geometric computing

6.2.1. Resultant of an equivariant polynomial system with respect to the symmetric group

Participants: Laurent Busé, Anna Karasoulou.

Given a system of n homogeneous polynomials in n variables which is equivariant with respect to the canonical actions of the symmetric group of n symbols on the variables and on the polynomials, it is proved that its resultant can be decomposed into a product of several smaller resultants that are given in terms of some divided differences. As an application, we obtain a decomposition formula for the discriminant of a multivariate homogeneous symmetric polynomial.

This work is submitted for publication [14].

6.2.2. Delaunay Mesh Generation of NURBS Surfaces

Participant: Laurent Busé.

We introduce a method for isotropic triangle meshing of NURBS surfaces. Based on Delaunay filtering and refinement, our approach departs from previous work by meshing in embedding space instead of parametric space. The meshing engine relies upon a novel line/surface intersection test, based on the matrix-based implicit representation of NURBS surfaces and numerical methods in linear algebra such as singular value and eigenvalue decompositions. A careful treatment of degenerate cases makes our approach robust to intersection points with multiple pre-images. In addition to ensure both approximation accuracy and mesh quality, our approach is seamless as it does not depend on the initial decomposition into NURBS patches, and is oblivious to the parameterization of the patches. Removing such dependencies provides us with a means to reliably mesh across patches with greater control over mesh sizing and shape of the elements.

This work was done in collaboration with Jingjing Shen and Neil Dodgson from Cambridge University and Pierre Alliez from TITANE.

6.2.3. Toric Border Basis

Participant: Bernard Mourrain.

In [11], we extend the theory and the algorithms of Border bases to systems of Laurent polynomial equations, defining “toric” roots. Instead of introducing new variables and new relations to saturate by the variable inverses, we propose a more efficient approach which works directly with the variables and their inverse. We show that the commutation relations and the inversion relations characterize toric border bases. We explicitly describe the first syzygy module associated to a toric border basis in terms of these relations. Finally, a new border basis algorithm for Laurent polynomials is described and a proof of its termination is given for zero-dimensional toric ideals.

Joint work with Philippe Trébuchet (LIP6 - UPMC).

6.2.4. Border Basis relaxation for polynomial optimization

Participants: Marta Abril-Bucero, Bernard Mourrain.

A relaxation method based on border basis reduction which improves the efficiency of Lasserre’s approach is proposed to compute the optimum of a polynomial function on a basic closed semi-algebraic set. A new stopping criterion is given to detect when the relaxation sequence reaches the minimum, using a sparse flat extension criterion. We also provide a new algorithm to reconstruct a finite sum of weighted Dirac measures from a truncated sequence of moments, which can be applied to other sparse reconstruction problems. As an application, we obtain a new algorithm to compute zero-dimensional minimizer ideals and the minimizer points or zero-dimensional G-radical ideals. Experimentations show the impact of this new method on significant benchmarks. See [12].

6.2.5. Flat extensions in $*$-algebra

Participant: Bernard Mourrain.

The objective of this work is to develop a flat extension characterization on moment matrices in the non-commutative case. We give a flat extension theorem for positive linear functionals on $_*$-algebras. The theorem is applied to truncated moment problems on cylinder sets, on matrices of polynomials and on enveloping algebras of Lie algebras. See [17].

Joint work with Konrad Schmüdgen, University of Leipzig, Germany.

6.3. Symbolic-Numeric Analysis

6.3.1. Cubatures, and related problems, with symmetry

Participants: Mathieu Collowald, Evelyne Hubert.

We address the computation of cubature formulae as a moment problem. Symmetry by finite groups arise naturally for cubatures. We developed the algebraic results to use the symmetry in order to reduce the number of parameters and the size of the matrices involved in the flat extension.

6.3.2. Quantitative Equidistribution for the Solutions of Systems of Sparse Polynomial Equations

Participant: André Galligo.

For a system of Laurent polynomials $f_1, ..., f_n \in \mathbb{C}[x_1^{\pm 1}, ..., x_n^{\pm 1}]$ whose coefficients are not too big with respect to its directional resultants, we show in [6] that the solutions in the algebraic torus $(\mathbb{C}^*)^n$ of the system of equations $f_1 = \cdots = f_n = 0$, are approximately equidistributed near the unit polycircle. This generalizes to the multivariate case a classical result due to Erdös and Turán on the distribution of the arguments of the roots of a univariate polynomial. We apply this result to bound the number of real roots of a system of Laurent polynomials, and to study the asymptotic distribution of the roots of systems of Laurent polynomials over $\mathbb{Z}$ and of random systems of Laurent polynomials over $\mathbb{C}$.

Joint work with Carlos D'Andrea (DM-UBA - Departamento de Matemática, Spain), Martin Sombra (ICREA & Universitat de Barcelona, Spain).

GALEN Project-Team

6. New Results

6.1. Highlights of the Year

- Handbook of Biomedical Imaging: Methodologies and Clinical Research [38] - co-edited from Nikos Paragios, James Duncan and Nicholas Ayache - has been published from Springer Publishing house.
- Nikos Paragios was admitted as a senior fellow at the Insitut Universitaire de France and has been awarded an IBM Faculty award. He has also been one of the plenary invited lecturers at the IARP International Conference in Pattern Recognition (ICPR'2015, Stockholm).

BEST PAPER AWARD :

[26] **Sparsity Techniques in Medical Imaging (STMI)**. M. MISYRLIS, A. KONOVA, M. BLASCHKO, J. HONORIO, N. ALIA-KLEIN, R. GOLDSTEIN, D. SAMARAS.

6.2. Rounding-based Moves for Metric Labeling

Paticipants: M. Pawan Kumar

Metric labeling is an important special case of energy minimizaton in Markov random fields. While the best known polynomial-time algorithm for the problem is the linear programming (LP) relaxation, in practice it is slow to solve it. In [25], we introduced a new family of efficient move-making algorithms for metric labeling. These algorithms mimic the rounding procedues used for converting a fractional LP solution to a feasible integral solution. Our algorithms provide a matching theoretical guarantee to the LP relaxation, while requiring significantly less computational time.

6.3. Optimizing Average Precision

Paticipants: Puneet Kumar Dokania, Aseem Behl, Pritish Mohapatra, C.V. Jawahar, M. Pawan Kumar

Average precision (AP) is one of the most commonly used measures for ranking. However, due to the inefficiency of optimizing it during learning, a common approach is to use surrogate loss functions such as 0-1 loss. In [27], we proposed a new optimization algorithm for AP-SVM that allows training in a similar time to binary SVM. In [23], we extended the AP-SVM framework to score the samples according to high-order information, as opposed to simple first-order information used in prior work. Finally, in [19], we proposed a novel latent AP-SVM formulation that allows learning from weakly supervised datasets. The advantage of learning with high order and missing information is demonstrated on challenging computer vision problems such as action classification and object detection using standard benchmark datasets.

6.4. Discriminative Training of Deformable Contour Models

Paticipants: Haithem Boussaid, Iasonas Kokkinos and Nikos Paragios

Deformable Contour Models (DCMs) are a main workhorse for medical image analysis - but are not commonly studied from a machine learning perspective. In [21], [20] we haved proposed an integrated machine learning and optimization framework to deploy DCMs in medical image analysis.

Our technical contributions are two-fold: firstly, we use an efficient decomposition-coordination algorithm to solve the optimization problems resulting from Loopy DCMs, by means of the Alternating Direction Method of Multipliers (ADMM); this yields substantially faster convergence than plain Dual Decomposition-based methods.

Secondly, we use structured prediction to exploit loss functions that better reflect the performance criteria used in medical image segmentation. By using the mean contour distance (MCD) as a structured loss during training, we obtain clear test-time performance gains.

We demonstrate the merits of exact and efficient inference with rich, structured models in a large X-Ray image segmentation benchmark, where we obtain systematic improvements over the current state-of-the-art.

6.5. Improved Deformable Part Models for Object Detection

Paticipants: Iasonas Kokkinos, Stavros Tsogkas, Eduard Trulls, Pierre-Andre Savalle, George Papandreou.

In [30] and [36] we have worked on improving the classification accuracy of Deformable Part Models (DPMs) for object detection in two distinct manners. Firstly, in [30] we propose a technique to combine bottom-up segmentation, coming in the form of SLIC superpixels, with sliding window DPM detectors. The merit of our approach lies in 'cleaning up' the low- level features by exploiting the spatial support indicated by segmentation. - tion, for both the root and part filters of DPMs. We use these masks to construct enhanced, background- invariant features to train DPMs. We test our approach on the PASCAL VOC 2007, outperforming the standard DPM in 17 out of 20 classes, yielding an average increase of 1.7AP. Additionally, we demonstrate the robustness of this ap- proach, extending it to dense SIFT descriptors for large dis- placement optical flow.

Secondly, in [36] we have explored the potential of convolutional neural networks as feature extractors for detection with DPMs. In particular, we substitute the Histogram-of-Gradient features of DPMs with Convolutional Neural Network (CNN) features, and demonstrate that we thereby obtain a substantial boost in performance (+14.5 mAP) when compared to the baseline HOG-based models. Some more recent extensions to this work are included in [41] where we explore the potential of explicit scale and aspect ratio search in the context of sliding window detection with CNNs.

6.6. Fine-Grained models of objects and texture

Paticipants: Iasonas Kokkinos, Matthew Blaschko, Stavros Tsogkas, Andrea Vedaldi, Mircea Cimpol, Subhransu Maji, Ross Girshick, Juho Kannala, Esa Rahtu, David Weiss, Ben Taskar, Karen Simonyan.

In [31] and [22] we explore methods for the fine-grained understanding of objects and textures, respectively.

In [22] we introduce a texture dataset that is accompanied by descriptions that capture the essence of the textures in terms of attributes. We explore a broad range of classification techniques for these texture attributes and demonstrate that the learned classifiers help improve generic texture recognition methods.

In [31] we introduce a large-scale dataset of airplanes that is accompanied by thorough human annotations at different levels: airplane types, segment lineouts, attributes, and part descriptions are provided for more than 7000 airplane images. We explore the potential of these rich annotations for the task of constructed fine-grained image descriptions using discriminative training techniques on top of standard image representations (Histogram-of-gradient features).

6.7. Large Scale Video Segmentation

Paticipants: Matthew Blaschko

Spatio-temporal cues are powerful sources of information for segmentation in videos. In [24] we present an efficient and simple technique for spatio-temporal segmentation that is based on a low-rank spectral clustering algorithm. The complexity of graph-based spatio-temporal segmentation is dominated by the size of the graph, which is proportional to the number of pixels in a video sequence. In contrast to other works, we avoid oversegmenting the images into super-pixels and instead generalize a simple graph based image segmentation. Our graph construction encodes appearance and motion information with temporal links based on optical flow. For large scale data sets naïve graph construction is computationally and memory intensive, and has only been achieved previously using a high power compute cluster. We make feasible for the first time large scale graph-based spatio-temporal segmentation on a single core by exploiting the sparsity structure of the problem and a low rank factorization that has strong approximation guarantees.

6.8. Higher Order Graph Matching

Paticipants: Chaohui Wang, Dimitris Samaras, Nikos Paragios

In [42] a generic framework for sparse and dense graph/3D surface matching has been introduced. The framework is endowed with a novel mathematical formulation regarding the matching process along with a novel deformation model. It exploits the power of invariance of higher order clique potentials and through a low to high resolution approach determines optimal correspondences between two sets of 3D points while taking advantage of Mobius tranformation to measure local similarity of shapes/graphs/surfaces. Graph matching of objects undergoing non-rigid deformations along with temporal 3D surface tracking demonstrated the potentials of our method. Inference is solved through an efficient dual decomposition scema.

6.9. Inference of Procedural Grammars from Images

Paticipants: Nikos Paragios

Grammar-like representations are powerful modeling and inference tools in computational vision. In [39] a novel approach towards automatic inference of typology specific building grammars has been introduced. The central idea was to consider that such grammars could be derived through a bottom up approach of common sub-tree reasoning of derivation trees determined through parsing using elementary shape (binary split) grammars. Such an approach performs common subtree reduction within the entire training set and identifies meta-rules (corresponding to the same subtrees) which are then clustered together towards producing a compact, typology specific grammar. Promising results both in terms of grammar compactness as well as in terms of inference demonstrated the potentials of the method that could be used beyond the considered scoped.

6.10. Fully connected CRFs for blood vessel segmentation in retinal images

Paticipants: Matthew Blaschko, José Ignacio Orlando

In [28], we present a novel method for blood vessel segmentation in fundus images based on a discriminatively trained, fully connected conditional random field model. Retinal image analysis is greatly aided by blood vessel segmentation as the vessel structure may be considered both a key source of signal, e.g. in the diagnosis of diabetic retinopathy, or a nuisance, e.g. in the analysis of pigment epithelium or choroid related abnormalities. Blood vessel segmentation in fundus images has been considered extensively in the literature, but remains a challenge largely due to the desired structures being thin and elongated, a setting that performs particularly poorly using standard segmentation priors such as a Potts model or total variation.

6.11. Graph-based Segmentation

Paticipants: Sarah Parisot, Deepak Chittajallu, Ioannis Kakadiaris, Nikos Paragios

In [17] we revisited explicit contour-evolution segmentation methods driven from a graph-based shape prior. Prior knowledge through geometric constraints has been encoded to the model within pair-wise interactions between control points. The segmentation process was driven through an objective function seeking to move the control points towards image locations optimizing the expected visual properties of the organ while satisfying the prior geometric constraints being learned at training. In [18] we have proposed a mathematical formalism for automatic tumor segmentation which was taking advantage of conventional segmentation likelihoods and atlas-based segmentation methods. The central idea was to jointly deform and segment an atlas such that the tumor likelihoods are maximized once projected to the targeted image while relaxing the registration constraints in this area. Furthermore we have endowed to this framework explicit estimation of uncertainties allowing the dynamic sampling of the graph structure resulting on significant speed up of the process while producing quantitative means for the interpretation of the final result.

6.12. Multi-atlas Segmentation

Paticipants: Stavros Alchatzidis, Aristeidis Sotiras, Nikos Paragios

In [33] a novel approach that couples pair-wise deformable registration with multi-atlas segmentation using graphical models was proposed. The method exploits both spaces and seeks to determine the optimal solution which will create the best possible visual agreement between atlases and target image along with their label consistency. The approach optimizes the deformation models and the segmentation labels jointly through an interconnected graph allowing either to relax registration constraints when segmentation labels do indicate or the opposite. The joint optimization of both spaces allowing the "implicit" automatic selection of atlases and therefore improves significantly segmentation performance.

6.13. Higher Order Graph Training throuh Dual Decomposition and Max Margin Principles

Paticipants: Nikos Komodakis, Bo Xiang, Nikos Paragios

In [40] a novel framework based on the structure margin principle was introduced for training higher order graphical models. The idea was to reduce the training of a complex high-order MRF in the parallel training of a series of simple slave MRFs through a principled dual decomposition approach. The theoretical properties of the framework have been studied while the method has been experimentally tested using 2d/3d segmentation problems involving higher order geometric priors that are linear-invariant. The proposed formulation benefits from theoretical guarantees as it concerns performance, computational simplicity while being modular and scalable.

GALLIUM Project-Team

6. New Results

6.1. Formal verification of compilers and static analyzers

6.1.1. Formal verification of static analyzers based on abstract interpretation

Participants: Jacques-Henri Jourdan, Xavier Leroy, Sandrine Blazy [EPI Celtique], Vincent Laporte [EPI Celtique], David Pichardie [EPI Celtique], Sylvain Boulmé [Grenoble INP, VERIMAG], Alexis Fouilhe [Université Joseph Fourier de Grenoble, VERIMAG], Michaël Périn [Université Joseph Fourier de Grenoble, VERIMAG].

In the context of the ANR Verasco project, we are investigating the formal specification and verification in Coq of a realistic static analyzer based on abstract interpretation. This static analyzer handles a large subset of the C language (the same subset as the CompCert compiler, minus recursion and dynamic allocation); supports a combination of abstract domains, including relational domains; and should produce usable alarms. The long-term goal is to obtain a static analyzer that can be used to prove safety properties of real-world embedded C codes.

This year, Jacques-Henri Jourdan continued the developpment of this static analyzer. He finished the proof of correctness of the abstract interpreter, using an axiomatic semantics for the C#minor intermediate language to decompose this proof in two manageable halves. He improved the precision and performance of the abstract iterator and of numerical abstract domains. He designed and verified a symbolic domain that helps analyzing sequential Boolean operators such as `&&` and `||` that are encoded as Boolean variables and conditional constructs in the C#minor intermediate language. As a more flexible alternative to reduced products of domains, Jacques-Henri Jourdan designed, implemented and proved correct a communication system between numerical abstract domains, based on communication channels and inspired by Astrée [56].

In parallel, IRISA and VERIMAG, our academic partners on the Verasco project, contributed a verified abstract domain for memory states and pointer values (Vincent Laporte, Sandrine Blazy, and David Pichardie) and a polyhedric abstract domain for linear numerical inequalities (Alexis Fouilhe, Sylvain Boulmé, Michaël Périn) that uses validation a posteriori. Those various components were brought together by Jacques-Henri Jourdan and Vincent Laporte, resulting in an executable static analyzer.

The overall architecture and specification of Verasco is described in a paper [29] accepted for presentation at the forthcoming POPL 2015 conference.

6.1.2. The CompCert formally-verified compiler

Participants: Xavier Leroy, Jacques-Henri Jourdan.

In the context of our work on compiler verification (see section 3.3.1), since 2005 we have been developing and formally verifying a moderately-optimizing compiler for a large subset of the C programming language, generating assembly code for the PowerPC, ARM, and x86 architectures [5]. This compiler comprises a back-end part, translating the Cminor intermediate language to PowerPC assembly and reusable for source languages other than C [4], and a front-end translating the CompCert C subset of C to Cminor. The compiler is mostly written within the specification language of the Coq proof assistant, from which Coq's extraction facility generates executable Caml code. The compiler comes with a 50000-line, machine-checked Coq proof of semantic preservation establishing that the generated assembly code executes exactly as prescribed by the semantics of the source C program.

This year, we improved the CompCert C compiler in several directions:

- The parser, previously compiled to unverified OCaml code, was replaced by a parser compiled to Coq code and validated *a posteriori* by a validator written and proved sound in Coq. This validation step, performed when the CompCert compiler is compiled, provides a formal proof that the parser recognizes exactly the language described by the source grammar. This approach builds on the earlier work by Jacques-Henri Jourdan, François Pottier and Xavier Leroy on verified validation of $LR(1)$ parsers [60]. Jacques-Henri Jourdan succeeded in scaling this approach all the way up to the full ISO C99 grammar plus some extensions.
- Two new static analyses, value analysis and neededness analysis, were added to the CompCert back-end. As described in section 6.1.3 below, the results of these analyses enable more aggressive optimizations over the RTL intermediate form.
- As part of the work on formalizing floating-point arithmetic (see section 6.1.4 below), the semantics and compilation of floating-point arithmetic in CompCert was revised to handle single-precision floating-point numbers as first-class values, instead of systematically converting them to double precision before arithmetic. This increases the efficiency and compactness of the code generated for applications that make heavy use of single precision.
- Previously, the CompCert back-end compiler was assuming a partitioned register set from the target architecture, where integer registers always contain 32-bit integers or pointers, and floating-point registers always contain double-precision FP numbers. This convention on register uses simplified the verification of CompCert, but became untenable with the introduction of single-precision FP numbers as first-class values: FP registers can now hold either single- or double-precision FP numbers. Xavier Leroy rearchitected the register allocator and the stack materialization passes of CompCert, along with their soundness proofs, to lift this limitation on register uses. Besides mixtures of single- and double-precision FP numbers, this new architecture makes it possible to support future target processors with a unified register set, such as the SPE variant of PowerPC.
- We added support for several features of ISO C99 that were not handled previously: designated initializers, compound literals, `switch` statements where the `default` case is not the last case, `switch` statements over arguments of 64-bit integer type, and incomplete arrays as the last member of a `struct`. Also, variable-argument functions and the `<stdarg.h>` standard include are now optionally supported, but their implementation is neither specified nor verified.
- The ARM back-end was extended with support for the EABI-HF calling conventions (passing FP arguments and results in FP registers instead of integer registers) and with generation of Thumb2 instructions. Thumb2 is an alternate instruction set and instruction encoding for the ARM architecture that results in more compact machine code (up to 30% reduction in code size on our tests).

We released three versions of CompCert, integrating these enhancements: version 2.2 in February 2014, version 2.3 in April, and version 2.4 in September.

In June 2014, Inria signed a licence agreement with AbsInt Angewandte Informatik GmbH, a software publisher based in Saarbrucken, Germany, to market and provide support for the CompCert formally-verified C compiler. AbsInt will extend CompCert to improve its usability in the critical embedded software market, and also provide long-term maintenance as required in this market.

6.1.3. Value analysis and neededness analysis in CompCert

Participant: Xavier Leroy.

Xavier Leroy designed, implemented, and proved sound two new static analyses over the RTL intermediate representation of CompCert. Both analyses are of the intraprocedural dataflow kind.

- Value analysis is a forward analysis that tracks points-to information for pointers, constantness information for integer and FP numbers, and variation intervals for integer numbers, using intervals of the form $[0, 2^n)$ and $[-2^n, 2^n)$. This value analysis extends and generalizes CompCert's earlier constant analysis as well as the points-to analysis of Robert and Leroy [68]. In particular, it tracks both the values of variables and the contents of memory locations, and it can take advantage of points-to information to show that function-local memory does not escape the scope of the function.

- Neededness analysis is a backward analysis that tracks which memory locations and which bits of the values of integer variables may be used later in a function, and which memory locations and integer bits are "dead", i.e. never used later. This analysis extends CompCert's earlier liveness analysis to memory locations and to individual bits of integer values.

Compared with the static analyses developed as part of Verasco (section 6.1.1), value analysis is much less precise: every function is analyzed independently of its call sites, relations between variables are not tracked, and even interval analysis is coarser (owing to CompCert's lack of support for widened fixpoint iteration). However, CompCert's static analyses are much cheaper than Verasco's, and scale well to large source codes, making it possible to perform them at every compilation run.

Xavier Leroy then modified CompCert's back-end optimizations to take advantage of the results of the two new static analyses, thus improving performance of the generated code:

- Common subexpression elimination (CSE) takes advantage of non-aliasing information provided by value analysis to eliminate redundant memory loads more aggressively.
- Many more integer casts (type conversions) and bit masking operations are discovered to be redundant and eliminated.
- Memory stores and block copy operations that become useless after constant propagation and CSE can now be eliminated entirely.

6.1.4. Verified compilation of floating-point arithmetic

Participants: Sylvie Boldo [EPI Toccata], Jacques-Henri Jourdan, Xavier Leroy, Guillaume Melquiond [EPI Toccata].

In 2012, we replaced the axiomatization of floating-point numbers and arithmetic operations used in early versions of CompCert by a fully-formal Coq development, building on the Coq formalization of IEEE-754 arithmetic provided by the Flocq library of Sylvie Boldo and Guillaume Melquiond. This verification of FP arithmetic and of its compilation was further improved in 2013 with respect to the treatment of "Not a Number" special values.

This year, Guillaume Melquiond improved the algorithmic efficiency of some of the executable FP operations provided by Flocq. Xavier Leroy generalized the theorems over FP arithmetic used in CompCert's soundness proof so that these theorems apply both to single- and double-precision FP numbers. Jacques-Henri Jourdan and Xavier Leroy proved additional theorems concerning conversions between integers and FP numbers.

A journal paper describing this 3-year work on correct compilation of floating-point arithmetic was accepted for publication at Journal of Automated Reasoning [14].

6.1.5. Verified JIT compilation of Coq

Participants: Maxime Dénès, Xavier Leroy.

Evaluation of terms from Gallina, the functional language embedded within Coq, plays a crucial role in the performance of proof checking or execution of verified programs, and the trust one can put in them. Today, Coq provides various evaluation mechanisms, some internal, in the kernel, others external, via extraction to OCaml or Haskell. However, we believe that the specific performance trade-offs and the delicate issues of trust are still calling for a better, more adapted, treatment.

That is why we started in October this year the Coqonut project, whose objective is to develop and formally verify an efficient, compiled implementation of Coq reductions. As a first step, we wrote an unverified prototype in OCaml producing x86-64 machine code using a monadic intermediate form. We started to port it to Coq and to specify the semantics of the source, target and intermediate languages.

6.2. Language design and type systems

6.2.1. The Mezzo programming language

Participants: Thibaut Balabonski, François Pottier, Jonathan Protzenko.

Mezzo is a programming language proposal whose untyped foundation is very much like OCaml (i.e., it is equipped with higher-order functions, algebraic data structures, mutable state, and shared-memory concurrency) and whose type system offers flexible means of describing ownership policies and controlling side effects.

In 2013 and early 2014, Thibaut Balabonski and François Pottier re-worked the machine-checked proof of type soundness for Mezzo. They developed a version of the proof which includes concurrency and dynamically-allocated locks, and showed that well-typed programs do not crash and are data-race free. This work was presented by François Pottier at FLOPS 2014 [24]. The proof was then extended with a novel and simplified account of adoption and abandon, a mechanism for combining the static ownership discipline with runtime ownership tests. A comprehensive paper, which contains both a tutorial introduction to Mezzo and a description of its formal definition and proof, was submitted to TOPLAS.

Minor modifications were carried out by Jonathan Protzenko in the implementation. A version of Mezzo that runs in a Web browser was developed and uploaded online, so that curious readers can play with the language without installing the software locally.

Jonathan Protzenko wrote his Ph.D. dissertation [12], which describes the design of Mezzo and the implementation of the Mezzo type-checker. He defended on September 29, 2014.

Web site: http://protz.github.io/mezzo/

6.2.2. System F with coercion constraints

Participants: Julien Cretin [Trust In Soft], Didier Rémy, Gabriel Scherer.

Expressive type systems often allow non trivial conversions between types, which may lead to complex, challenging, and sometimes ad hoc type systems. Such examples are the extension of System F with type equalities to model GADTs and type families of Haskell, or the extension of System F with explicit contracts. A useful technique to simplify the meta-theoretical study of such systems is to view type conversions as *coercions* inside terms.

Following a general approach based on System F, Julien Cretin and Didier Rémy earlier introduced a language of *explicit coercions* enabling abstraction over coercions and viewing all type transformations as explicit coercions [57]. To ensure that coercions are erasable, *i.e.*, that they only decorate terms without altering their reduction, they are restricted to those that are parametric in either their domain or codomain. Despite this restriction, this language already subsumed many extensions of System F, including bounded polymorphism, instance-bounded polymorphism, and η-conversions—but not subtyping constraints.

To lift this restriction, Julien Crétin and Didier Rémy proposed a new approach were coercions are left implicit. Technically, we extended System F with a rich language of propositions containing a first-order logic, a coinduction mechanism, consistency assertions, and coercions (which are thus just a particular form of propositions); we then introduced a type-level language using kinds to classify types, and constrained kinds to restrict kinds to types satisfying a proposition. Abstraction over types of a constrained kind amounts to abstraction over arbitrary propositions, including coercions.

By default, type abstraction should be erasable, which is the case when kinds of abstract type variables are inhabited—we say that such abstractions are consistent. Still, inconsistent type abstractions are also useful, for instance, to model GADTs. We provide them as a different construct, since they are not erasable, as they must delay reduction of subterms that depend on them. This introduces a form of weak reduction in a language with full reduction, which is a known source of difficulties: although the language remains sound, we loose the subject reduction property. This work has been described in [28] and is part of Julien Cretin's PhD dissertation [11] defended in January 2014; a simpler, core subset is also described in [45].

Recently, Gabriel Scherer and Didier Rémy introduced *assumption hiding* [32], [50] to restore confluence when mixing full and weak reductions and provide a continuum between consistent and inconsistent abstraction. Assumption hiding allows a fine-grained control of dependencies between computations and the logical hypotheses they depend on; although studied for a language of coercions, the solution is more general and should be applicable to any language with abstraction over propositions that are left implicit, either for the user's convenience in a surface language or because they have been erased prior to computation in an internal language.

6.2.3. Singleton types for code inference

Participants: Gabriel Scherer, Didier Rémy.

We continued working on singleton types for code inference. If we can prove that a type contains, in a suitably restricted pure lambda-calculus, a unique inhabitant modulo program equivalence, the compiler can infer the code of this inhabitant. This opens the way to type-directed description of boilerplate code, through type inference of finer-grained type annotations. A decision algorithm for the simply-typed lambda-calculus is still work-in-progress. We presented at the TYPES'14 conference [42] our general approach to such decision procedures, and obtained an independent counting result for intuitionistic logic [52] that demonstrates the finiteness of the search space.

6.2.4. Generic programming with ornaments

Participants: Pierre-Évariste Dagand, Didier Rémy, Thomas Williams.

Since their first introduction in ML, datatypes have evolved: besides providing an organizing *structure* for computation, they are now offering more *control* over what is a valid result. GADTs, which are now part of the OCaml language, offer such a mechanism: ML programmers can express fine-grained, logical invariants of their datastructures. Programmers thus strive to express the correctness of their programs in the types: a well-typed program is correct by construction. However, these carefully crafted datatypes are a threat to any library design: the same data-*structure* is used for many logically incompatible purposes. To address this issue, McBride developed *ornaments*. It defines conditions under which a new datatype definition can be described as an ornament of another one, typically when they both share the same inductive definition scheme. For example, lists can be described as the ornament of the Church encoding of natural numbers. Once a close correspondence between a datatype and its ornament has been established, certain kinds of operations on the original datatype can be automatically lifted to its ornament.

To account for whole-program transformations, we developed a type-theoretic presentation of *functional ornament* [17] as a generalization of ornaments to functions. This work built up on a type-theoretic *universe of datatypes*, a first-class description of inductive types within the type theory itself. Such a presentation allowed us to analyze and compute over datatypes in a transparent manner. Upon this foundation, we formalized the concept of functional ornament by another type-theoretic universe construction. Based on this universe, we established the connection between a base function (such as addition and subtraction) and its ornamented version (such as, respectively, the concatenation of lists and the deletion of a prefix). We also provided support for driving the computer into semi-automatically lifting programs: we showed how addition over natural numbers could be incrementally evolved into concatenation of lists.

Besides the theoretical aspects, we have also tackled the practical question of offering ornaments in an ML setting [33]. Our goal was to extend core ML with support for ornaments so as to enable semi-automatic program transformation and fully-automatic code refactoring. We thus treated the purely syntactic aspects, providing a concrete syntax for describing ornaments of datatypes and specifying the lifting of functions. Such lifting specifications allow the user to declaratively instruct the system to, for example, lift addition of numbers to concatenation of lists. We gave an algorithm that, given a lifting specification, performs the actual program transformation from the bare types to the desired, ornamented types. This work has been evaluated by a prototype implementation in which we demonstrated a few typical use-cases for the semi-automatic lifting of programs.

Having demonstrated the benefits of ornaments in ML, it has been tempting to offer ornaments as first-class citizens in a programming language. Doing so, we wished to rationalize the lifting of programs as an elaboration process within a well-defined, formal system. To describe the liftings, one would like to specify only the local transformations that are applied to the original program. Designing such a language of *patches* and formalizing its elaboration has been the focus of our recent efforts.

6.2.5. Constraints as computations

Participant: François Pottier.

Hindley-Milner type inference–the problem of determining whether an ML program is well-typed–is well-understood, and can be elegantly explained and implemented in two phases, namely constraint generation and constraint solving. In contrast, elaboration–the task of constructing an explicitly-typed representation of the program–seems to have received relatively little attention in the literature, and did not until now enjoy a modular constraint-based presentation. François Pottier proposed such a presentation, which views constraints as computations and equips them with the structure of an applicative functor. This work was presented as a "functional pearl" at ICFP 2014 [31]. The code, in the form of a re-usable library, is available online.

6.2.6. Equivalence and normalization of lambda-terms with sums

Participants: Gabriel Scherer, Guillaume Munch-Maccagnoni [Université 13, LIPN lab].

Determining uniqueness of inhabitants requires a good understanding of program equivalence in presence of sum types. In yet-unpublished work, Gabriel Scherer worked on the correspondence between two existing normalization techniques, one coming from the focusing community [54] and the other using direct lambda-term rewriting [63]. A collaboration with Guillaume Munch-Maccagnoni has also started this year, whose purpose is to present normalization procedures for sums using System L, a rich, untyped syntax of terms (or abstract machines) for the sequent calculus.

6.2.7. Computational interpretation of realizability

Participants: Pierre-Évariste Dagand, Gabriel Scherer.

We are trying to better understand the computational behavior of semantic normalization techniques such as a realizability and logical relation models. As a very first step, we inspected the computational meaning of a normalization proof by realizability for the simply-typed lambda-calculus. It corresponds to an evaluation function; the evaluation order for each logical connective is determined by the definition of the sets of truth and value witnesses. This preliminary work is to be presented at JFLA 2015 [35].

6.3. Shared-memory parallelism

6.3.1. Algorithms and data structures for parallel computing

Participants: Umut Acar, Arthur Charguéraud [EPI Toccata], Mike Rainey.

The ERC Deepsea project, with principal investigator Umut Acar, started in June 2013 and is hosted by the Gallium team. This project aims at developing techniques for parallel and self-adjusting computations in the context of shared-memory multiprocessors (i.e., multicore platforms). The project is continuing work that began at Max Planck Institute for Software Systems between 2010 and 2013. As part of this project, we are developing a C++ library, called PASL, for programming parallel computations at a high level of abstraction. We use this library to evaluate new algorithms and data structures. We obtained two major results this year.

The first result is a sequence data structure that provides amortized constant-time access at the two ends, and logarithmic time concatenation and splitting at arbitrary positions. These operations are essential for programming efficient computation in the fork-join model. Compared with prior work, this novel sequence data structure achieves excellent constant factors, allowing it to be used as a replacement for traditional, non-splittable sequence data structures. This data structure, called *chunked sequence* due to its use of chunks (fixed-capacity arrays), has been implemented both in C++ and in OCaml, and shown competitive with state-of-the art sequence data structures that do not support split and concatenation operations. This work is described in a paper published at ESA [22].

A second main result is the development of fast and robust parallel graph traversal algorithms, more precisely for parallel BFS and parallel DFS. The new algorithms leverage the aformentioned sequence data structure for representing the set of edges remaining to be visited. In particular, it uses the split operation for balancing the edges among the several processors involved in the computation. Compared with prior work, these new algorithms are designed to be efficient not just for particular classes of graphs, but for all input graphs. This work has not yet been published, however it is described in details in a technical report [46].

6.3.2. Weak memory models

Participants: Luc Maranget, Jacques-Pascal Deplaix, Jade Alglave [University College London, then Microsoft Research, Cambridge].

Modern multi-core and multi-processor computers do not follow the intuitive "Sequential Consistency" model that would define a concurrent execution as the interleaving of the execution of its constituting threads and that would command instantaneous writes to the shared memory. This situation is due both to in-core optimisations such as speculative and out-of-order execution of instructions, and to the presence of sophisticated (and cooperating) caching devices between processors and memory.

In the last few years, Luc Maranget took part in an international research effort to define the semantics of the computers of the multi-core era. This research effort relies both on formal methods for defining the models and on intensive experiments for validating the models. Joint work with, amongst others, Jade Alglave (now at Microsoft Research, Cambridge), Peter Sewell (University of Cambridge) and Susmit Sarkar (University of St. Andrews) achieved several significant results, including two semantics for the IBM Power and ARM memory models: one of the operational kind [70] and the other of the axiomatic kind [64]. In particular, Luc Maranget is the main developer of the **diy** tool suite (see section 5.3). Luc Maranget also performs most of the experiments involved.

In 2014 we produced a new model for Power/ARM. The new model is simpler than the previous ones, in the sense that it is based on fewer mathematical objects and can be simulated more efficiently than the previous models. The new **herd** simulator (part of **diy** tool suite) is in fact a generic simulator, whose central component is an interpreter for a domain-specific language. More precisely, memory models are described in a simple language that defines relations by means of a few operators such as concatenation, transitive closure, fixpoint, etc., and performs validity checks on relations such as acyclicity. The Power/ARM model consists of about 50 lines of this specific language. This work, with additional material, including in-depth testing of ARM devices and data-mining of potential concurrency bugs in a huge code base, was published in the journal *Transaction on Programming Languages and Systems* [13] and selected for presentation at the PLDI conference [23]. Luc Maranget gave this presentation.

In the same research theme, Luc Maranget supervised the internship of Jacques-Pascal Deplaix (EPITECH), from Oct. 2013 to May 2014. Jacques-Pascal extended **litmus**, our tool to run tests on hardware. **litmus** now accepts test written in C; we can now perform the conformance testing of C compilers and machines with respect to the C11/C++11 standard. Namely, Mark Batty (University of Cambridge), under the supervision of Jade Alglave, wrote a **herd** model for this standard. The new **litmus** also proves useful to run tests that exploit some machine idiosyncrasies, when our **litmus** assembly implementation does not handle them.

As a part of the **litmus** infrastructure, Luc Maranget designed a synchronisation barrier primitive by simplifying the sense synchronisation barrier published by Maurice Herlily and Nir Shavit in their textbook [58]. He co-authored a JFLA article [34], that presents this primitive and proves it correct automatically by the means of the **cubicle** tool developed under the supervision of Sylvain Conchon (team Toccata, Inria Saclay).

6.4. The OCaml language and system

6.4.1. The OCaml system

Participants: Damien Doligez, Alain Frisch [Lexifi SAS], Jacques Garrigue [University of Nagoya], Fabrice Le Fessant, Xavier Leroy, Luc Maranget, Gabriel Scherer, Mark Shinwell [Jane Street], Leo White [OCaml Labs, Cambridge University], Jeremy Yallop [OCaml Labs, Cambridge University].

This year, we released versions 4.02.0 and 4.02.1 of the OCaml system. Release 4.02.0 is a major release that fixes about 60 bugs and introduces 70 new features suggested by users. Damien Doligez acted as release manager for both versions.

OCaml 4.02.0 introduces a large number of major innovations:

- Extension points: a uniform syntax for adding attributes and extensions in OCaml source code: most external preprocessors can now extend the language without need to extend the syntax and reimplement the parser.
- Improvements to the module system: generative functors and module aliases facilitate the efficient handling of large code bases.
- Separation between text-like read-only strings and array-like read-write byte sequences. This makes OCaml programs safer and clearer.
- An extension to the pattern-matching syntax to catch exceptions gives a short, readable way to write some important code patterns.
- Extensible open datatypes generalize the exception type and make its features available for general programming.
- Several important optimizations were added or enhanced: constant propagation, common subexpression elimination, dead code elimination, optimization of pattern-matching on strings.
- A code generator for the new 64-bit ARM architecture “AArch64”.
- A safer and faster implementation of the printf function, based on the GADT feature introduced in OCaml 4.00.0.

This version has also seen a reduction in size: the Camlp4 and Labltk parts of the system are now independent systems. This makes them free to evolve on their own release schedules, and to widen their contributor communities beyond the core OCaml team.

OCaml 4.02.1 fixes a few bugs introduced in 4.02.0, along with 25 older bugs.

In parallel, we designed and experimented with several new features that are candidates for inclusion in the next major releases of OCaml:

- Ephemerons: a more powerful version of weak pointers.
- A parallel extension of the runtime system and associated language features that will let multi-threaded OCaml programs run in parallel on several CPU cores.
- Modular implicits: a typeclass-like extension that will make is easy to write generic code (*e.g.* print functions, comparison predicates, overloaded arithmetic operators, etc).
- “Inlined” records as constructor arguments, which will let the programmer select a packed representation for important data structures.
- Major improvements to the inlining optimization pass.
- Support for debugging native-code OCaml programs with GDB.

6.4.2. Namespaces for OCaml

Participants: Fabrice Le Fessant, Pierrick Couderc.

With the growth of the OCaml community and the ease of sharing code through OPAM, the new OCaml package manager, OCaml projects are using more and more external libraries. As a consequence, conflicts between module names of different libraries are now more likely for big projects, and the need for switching from the current flat namespace to a hierarchical namespace is now real.

We experimented with a prototype of OCaml where the namespaces used by a module are explicitely written in the OCaml module source header, to generate the environment in which the source is typed and compiled [39]. Namespaces are mapped on directories on the disk. This mechanism complements the recent addition of module aliases to OCaml, by providing extensibility at the namespace level, whereas it is absent at the module level, and solves also the problem of exact dependency analysis (the previous tool used for that purpose, ocamldep, provides only an approximation of the dependencies, computed on the syntax tree).

6.4.3. Memory profiling OCaml application

Participants: Fabrice Le Fessant, Çagdas Bozman [ENSTA ParisTech], Grégoire Henry [OCamlPro], Michel Mauny [ENSTA ParisTech].

Most modern languages make use of automatic memory management to discharge the programmer from the burden of allocating and releasing the chunks of memory used by the software. As a consequence, when an application exhibits an unexpected usage of memory, programmers need new tools to understand what is happening and how to solve such an issue. In OCaml, the compact representation of values, with almost no runtime type information, makes the design of such tools more complex.

We have experimented with three tools to profile the memory usage of real OCaml applications. The first tool saves snapshots of the heap after every garbage collection. Snapshots can then be analysed to display the evolution of memory usage, with detailed information on the types of values, where they were allocated and from where they are still reachable. A second tool updates counters at every garbage collection event, it complements the first tool by providing insight on the behavior of the minor heap, and the values that are promoted or not to the major heap. Finally, a third tool samples allocations and saves stacks of function calls at these samples.

These tools have been used on real applications (Alt-Ergo, an SMT solver, or Cumulus, an Ocsigen website), and allowed us to track down and fix memory problems with these applications, such as useless copies of data structures and memory leaks.

6.4.4. OPAM, the OCaml package manager

Participants: Fabrice Le Fessant, Roberto Di Cosmo [IRILL], Louis Gesbert [OCamlPro].

With the growth of the OCaml community, the need for sharing libraries between users has lead to the development of a new package manager, called OPAM. OPAM is based on Dose, a library developed by the Mancoosi team at IRILL, to provide a unified format, CUDF, to query external dependency solvers. The specific needs of OPAM have driven interesting research and improvements on the Dose library, that have consequently opened new opportunities for improvements in OPAM, for the benefit of both software.

We have for example experimented with the design of a specific language [37] to describe optimization criteria, when managing OPAM packages. Indeed, depending on the actions (installation, upgrade, removal), the user might want to reach very different configurations, requiring an expressive power that go far beyond what traditional package managers can express in their configuration options. For example, during installation, the user would probably see as little compilation as possible, whereas upgrading is supposed to move the configuration to the most up-to-date state, with as much compilation as needed.

We have also proposed a new paradigm: multi-switch constraints, to model *switches* used in OPAM to handle different versions of OCaml on the same computer [41]. We proposed this new paradigm as a way to solve multiple problems (cross-compilation, multi-switch packages, per-switch repositories and application-specific switches). However, we expect this new paradigm to challenge the scalability of the current CUDF solvers used by OPAM, and to require important changes and optimization in the Dose library.

6.5. Software specification and verification

6.5.1. Tools for TLA+

Participants: Damien Doligez, Jael Kriener, Leslie Lamport [Microsoft Research], Stephan Merz [EPI VeriDis], Tomer Libal [Microsoft Research-Inria Joint Centre], Hernán Vanzetto [Microsoft Research-Inria Joint Centre].

Damien Doligez is head of the “Tools for Proofs” team in the Microsoft-Inria Joint Centre. The aim of this team is to extend the TLA+ language with a formal language for hierarchical proofs, formalizing the ideas in [61], and to build tools for writing TLA+ specifications and mechanically checking the corresponding formal proofs.

This year, we released two versions of the TLA+ Proof System (TLAPS), the part of the TLA+ tools that handles mechanical checking of TLA+ proofs. This environment is described in [55].

These versions add the propositional temporal logic prover LS4 as a back-end, which allows TLAPS to deal with propositional temporal formulas. This relies on a technique called *coalescing* [40], which allows users to prove arbitrary safety properties, as well as some liveness properties, by translating them into the back-end prover's logic without increasing the complexity of the formulas.

Jael Kriener started a post-doc contract in December 2013, funded by the ADN4SE contract, and left in September 2014. She worked on the theory of temporal proofs in TLA+ and, in collaboration with CEA, on proving some properties of the PharOS real-time operating system.

Web sites:
http://research.microsoft.com/users/lamport/tla/tla.html
http://tla.msr-inria.inria.fr/tlaps

6.5.2. The Zenon automatic theorem prover

Participants: Damien Doligez, David Delahaye [CNAM], Pierre Halmagrand [Equipe DEDUCTEAM], Guillaume Bury [Equipe DEDUCTEAM], Olivier Hermant [Mines ParisTech].

Damien Doligez continued the development of Zenon, a tableau-based prover for first-order logic with equality and theory-specific extensions.

Pierre Halmagrand continued his thesis work, funded by ANR BWare, on integrating Deduction Modulo in Zenon, with emphasis on making it efficient for dealing with B set theory.

Guillaume Bury did an internship, also funded by ANR BWare. He implemented an extension of Zenon, based on the simplex method, to deal with arithmetic formulas.

6.5.3. Well-typed generic fuzzing for module interfaces

Participants: Thomas Braibant, Jonathan Protzenko, Gabriel Scherer.

Property-based testing generates arbitrary instances of inputs to check a given correctness predicate/property. Thomas Braibant proposed that, instead of a random generation function defined from the internals of one's data-structure, one could use the user-exposed interface to generate instances by composition of API calls. GADTs let us reflect/reify a typed API, and program a type-respecting exploration/testing directly in the host language. We developed a prototype library, Articheck, to experiment with this idea. This work was presented at the ML Workshop [38].

6.5.4. Depth-First Search and Strong Connectivity in Coq

Participant: François Pottier.

In 2002, Ingo Wegener published a short paper which sketches a proof of Kosaraju's linear-time algorithm for computing the strongly connected components of a directed graph. At the same time, Wegener's paper helps explain why the algorithm works, which, from a pedagogical standpoint, makes it quite valuable. In 2013 and 2014, François Pottier produced a machine-checked version of Wegener's proof, and wrote a precise informal account of it, which will be presented at JFLA 2015 [36].

6.5.5. Implementing hash-consed structures in Coq

Participants: Thomas Braibant, Jacques-Henri Jourdan, David Monniaux [CNRS, VERIMAG].

Hash-consing is a programming technique used to implement maximal sharing of immutable values in memory, keeping a single copy of semantically equivalent objects. Hash-consed data-structures give a unique identifier to each object, allowing fast hashing and comparisons of objects. This may lead to major improvements in execution time by itself, but it also makes it possible to do efficient memoization of computations.

Hash-consing and memoization are examples of imperative techniques that are of prime importance for performance, but are not easy to implement and prove correct using the purely functional language of a proof assistant such as Coq.

We published an article in Journal of Automated Reasoning [15], explaining our work on this subject during the last 3 years. We gave four different approaches for implementing hash-consed data-structures in Coq. Then, we performed an in-depth comparative study of how our "design patterns" for certified hash-consing fare on two real-scale examples: BDDs and lambda-terms.

GAMMA3 Project-Team

4. New Results

4.1. Serendipity and reduced elements

Participants: Paul Louis George [correspondant], Houman Borouchaki, Nicolas Barral.

We give a method to constructing Serendipity elements for quads and hexes with full symmetry properties and indicate the reading of their shape functions. We show that, since the degree 5, the Serendipity elements are no longer symmetric but we propose a method resulting in a Lagrange element of degree 5 with full symmetry properties after adding an adequate number of additional nodes.

On the other hand, we show how to guarantee the geometric validity of a given curved element (seen as a patch) of a mesh. This is achieved after writing the patch in a Bézier setting (Bernstein polynomials and control points). In addition, we discuss the case of patch derived from a transfinite interpolation and it is proved that only some of them are Serendipity elements indeed, we return to the same elements as above

We also give a method to constructing Lagrange Serendipity (or reduced) simplices with a detailed description of the triangles of degree 3 and 4. We indicate that higher order triangles are not candidate apart if we impose a restricted polynomial space. We show that a tetrahedron of degree 3 is a candidate while high order elements are not candidate even if a restriction in the polynomial space is considered. In addition, we propose a method for the validation of such elements, in a given mesh, where the validation means the positiveness of the jacobian.

A technical report have been published [29].

4.2. Validity of transfinite and Bézier-Serendipity patches

Participants: Paul Louis George [correspondant], Houman Borouchaki, Nicolas Barral.

We define generalized transfinite patches for quads and hexes with full symmetry properties. We give a way of constructing those patches by considering the Bézier setting using linear combinations of tensor-product patches of various degree. Those patches are exactly the Bézier-Serendipity patches recently introduced

ASsfor reduced quadrilateral patches, we introduce the so called "Bézier-Serendip" patches. After some recalls about standard Bézier patches, we propose a method to constructing those reduced patches. The corresponding Bernstein polynomials are written by means of linear combinations of the standard Bernstein polynomials. We give a full description of the patches of degree 2, 3, 4 and 5. Since degree 5, the location of the control points is no longer symmetric and to remedy this problem, we propose adding a number of control points which results in *extended* Bézier-Serendip patches. Those reduced patches are in the Bézier framework what the Serendipity elements are in the finite element framework.

A technical report and a paper have been published [30], [17].

4.3. Meshing Strategies and the Impact of Finite Element Quality on the Velocity Field in Fractured Media

Participants: Patrick Laug [correspondant], Géraldine Pichot.

For calculating flow in a fracture network, the mixed hybrid finite element (MHFE) method is a method of choice as it yields a symmetric, positive definite linear system. However, a drawback to this method is its sensitivity to bad aspect ratio elements. For poor-quality triangles, elementary matrices are ill-conditioned, and inconsistent velocity vectors are obtained by inverting these local matrices. In this work, different strategies have been proposed for better reconstruction of the velocity field [21].

4.4. Automatic Mesh Generation of Multiface Models on Multicore Processors

Participant: Patrick Laug [correspondant].

This work started in September 2014, as part of a sabbatical at Polytechnique Montréal. In a previous study, a parallel version of an indirect approach for meshing composite surfaces – also called multiface models – was developed. However, this methodology could be inefficient in practice, as the memory management of most existing CAD (computer aided design) systems use static global caches to save information. In our new approach, CAD queries are fully parallelized, using the Pirate library from Polytechnique Montréal. This library provides a set of C++ classes that implement STEP-compliant B-Rep geometric and topological entities, as well as classes to represent meshes and solutions. By modifying the data structures so that memory caches are local to each face of the geometric model, geometric primitives can efficiently be evaluated in parallel, and performance measurements show significant gains.

4.5. Applications du maillage et développements de méthodes avancées pour la cryptographie

Participants: Thomas Grosges [correspondant], Dominique Barchiesi, Michael François.

L'utilisation des nombres (pseudo)-aléatoires a pris une dimension importante ces dernières décennies. De nombreuses applications dans le domaine des télécommunications, de la cryptographie, des simulations numériques ou encore des jeux de hasard, ont contribué au développement et à l'usage de ces nombres. Les méthodes utilisées pour la génération de tels nombres (pseudo)-aléatoires proviennent de deux types de processus : physique et algorithmique. Ce projet de recherche a donc pour objectif principal le développement de nouveaux procédés de génération de clés de chiffrement, dits "exotiques", basés sur des processus physiques, multi-échelles, multi-domaines assurant un niveau élevé de sécurité. Deux classes de générateurs basés sur des principes de mesures physiques et des processus mathématiques ont été développé.

La première classe de générateurs exploite la réponse d'un système physique servant de source pour la génération des séquences aléatoires. Cette classe utilise aussi bien des résultats de simulation que des résultats de mesures interférométriques pour produire des séquences de nombres aléatoires. L'application du maillage adaptatif sert au contrôle de l'erreur sur la solution des champs physiques (simulés ou mesurés). A partir de ces cartes physiques, un maillage avec estimateur d'erreur sur l'entropie du système est appliqué. Celui-ci permet de redistribuer les positions spatiales des noeuds. L'étude (locale) de la réduction d'entropie des clés tout au long de la chaîne de création et l'étude (globale) de l'entropie de l'espace des clés générées sont réalisées à partir de tests statistiques.

La seconde classe de générateurs porte sur le développement de méthodes avancées et est basée sur l'exploitation de fonctions chaotiques en utilisant les sorties de ces fonctions comme indice de permutation sur un vecteur initial. Ce projet s'intéresse également aux systèmes de chiffrement pour la protection des données et deux algorithmes de chiffrement d'images utilisant des fonctions chaotiques sont développés et analysés. Ces Algorithmes utilisent un processus de permutation-substitution sur les bits de l'image originale. Une analyse statistique approfondie confirme la pertinence des cryptosystèmes développés.

4.6. Développement de méthodes avancées et maillages appliqués à l'étude de la nanomorphologie des nanotubes-fils en suspension liquide

Participants: Thomas Grosges [correspondant], Dominique Barchiesi, Abel Cherouat, Houman Borouchaki, Laurence Giraud-Moreau, Anis Chaari.

Ce projet de recherche (NANOMORPH) a pour objet principal le développement et la mise au point d'une instrumentation optique pour déterminer la distribution en tailles et le coefficient de forme de nanofils (NF) ou de nanotubes (NT) en suspension dans un écoulement. Au cours de ce projet, deux types de techniques optiques complémentaires sont développées. La première, basée sur la diffusion statique de la lumière, nécessite d'étudier au préalable la physico-chimie de la dispersion, la stabilisation et l'orientation des nanofils dans les milieux d'étude. La seconde méthode, basée sur une méthode opto-photothermique pulsée, nécessite en sus, la modélisation de l'interaction laser/nanofils, ainsi que l'étude des phénomènes multiphysiques induits par ce processus. L'implication de l'équipe-projet GAMMA3 concerne principalement la simulation multiphysique de l'interaction laser-nanofils et l'évolution temporelle des bulles et leurs formations. L'une des principales difficultés de ces problématiques est que la géométrie du domaine est variable (à la fois au sens géométrique et topologique). Ces simulations ne peuvent donc être réalisées que dans un schéma adaptatif de calcul nécessitant le remaillage tridimensionnel mobile, déformable avec topologie variable du domaine (formation et évolution des bulles au cours du temps et de l'espace).

4.7. Applications du maillage à des problèmes multi-physiques, développement de méthodes de résolutions avancées et modélisation électromagnetisme-thermique-mécanique à l'échelle mesoscopique

Participants: Dominique Barchiesi [correspondant], Abel Cherouat, Thomas Grosges, Houman Borouchaki, Laurence Giraud-Moreau, Sameh Kessentini, Anis Chaari, Fadhil Mezghani.

Le contrôle et l'adaptation du maillage lors de la résolution de problèmes couplés ou/et non linéaires reste un problème ouvert et fortement dépendant du type de couplage physique entre les EDP à résoudre. Notre objectif est de développer des modèles stables afin de calculer les dilatations induites par l'absorption d'énergie électromagnétique, par des structures matérielles inférieures au micron. Les structures étudiées sont en particulier des nanoparticules métalliques en condition de résonance plasmon. Dans ce cas, un maximum d'énergie absorbée est attendu, accompagné d'un maximum d'élévation de température et de dilatation. Il faut en particulier développer des modèles permettant de simuler le comportement multiphysique de particules de formes quelconques, pour une gamme de fréquences du laser d'éclairage assez étendue afin d'obtenir une étude spectroscopique de la température et de la dilatation. L'objectif intermédiaire est de pouvoir quantifier la dilatation en fonction de la puissance laser incidente. Le calcul doit donc être dimensionné et permettre finalement des applications dans les domaines des capteurs et de l'ingénierie biomédicale. En effet, ces nanoparticules métalliques sont utilisées à la fois pour le traitement des cancers superficiels par nécrose de tumeur sous éclairage adéquat, dans la fenêtres de transparence cellulaire. Déposées sur un substrat de verre, ces nanoparticules permettent de construire des capteurs utilisant la résonance plasmon pour être plus sensibles (voir projet européen *Nanoantenna* et l'activité génération de nombres aléatoires). Cependant, dans les deux cas, il est nécessaire, en environnement complexe de déterminer la température locale, voire la dilatation de ces nanoparticules, pouvant conduire à un désaccord du capteur, la résonance plasmon étant très sensible aux paramètres géométriques et matériels des nanostructures. Dans ce sens, l'étude permet d'aller plus loin que la "simple" interaction électromagnétique avec la matière du projet européen *Nanoantenna*.

Le travail de l'année 2014 a constitué en la poursuite de l'étude des spécificités de ce type de problème multiphysique pour des structures de forme simple et la mise en place de fonctions test, de référence, pour les développements de maillage adaptatifs pour les modèles multiphysiques éléments finis. Nous espérons pouvoir proposer un projet ANR couplant les points de vue microscopiques et macroscopiques dans les deux années qui viennent.

4.8. Visualization and modification of high-order curved meshes

Participants: Alexis Loyer, Adrien Loseille [correspondant].

During the partnership between Inria and Distene, a new visualization software has been designed. It address the typical operations that are required to quickly assess the newly algorithm developed in the team. In particular, interactive modifications of high-order curved mesh and hybrid meshes has been addressed. The software VIZIR is freely available at https://www.rocq.inria.fr/gamma/gamma/vizir/.

4.9. Mesh adaptive ALE numerical simulation

Participants: Frédéric Alauzet [correspondant], Nicolas Barral, Adrien Loseille.

Running highly accurate numerical simulations with moving geometries is still a challenge today due to their prohibitive cost in CPU time. Using anisotropic mesh adaptation is one way to drastically reduce the size of the problem and to reach the desired accuracy. Previously, we have developed an ALE formulation using mesh connectivity change in order to achieve any complex displacement. Then, this method has been coupled with the unsteady anisotropic mesh adaptation using the fixed-point algorithm. The key point of this work is the use of an ALE metric that takes into account the mesh motion in the metric field definition [24], [14].

4.10. Mesh adaptation for Navier-Stokes Equations

Participants: Frédéric Alauzet, Victorien Menier, Adrien Loseille [correspondant].

Adaptive simulations for Navier-Stokes equations require to propose accurate error estimates and design robust mesh adaptation algorithms (for boundary layers).

For error estimates, we design new estimates suited to accurately capture the speed profile in the boundary layers. For mesh adaptation, we design a new method to generate structured boundary layer meshes which are mandatory to accurately compute compressible flows a high Reynolds number (several millions). It couple the specification of the optimal boundary layer from the geometry boundary and moving mesh techniques to extrude the boundary layer in an already existing mesh. The main advantage of this approach is its robustness, *i.e.,* at each step of the algorithm we have always a valid mesh [25].

4.11. Adaptive multigrid strategies

Participants: Frédéric Alauzet [correspondant], Victorien Menier, Adrien Loseille.

Multigrid is a well known technique used to accelerate the convergence of linear system solutions. Using a multigrid strategy to solve non-linear problems improves the robustness and the convergence of each Newton step, the accelerating overall the whole process. In particular, larger time step can be considered. This of main importance when solving turbulent Navier-Stokes equations on complex geometries. First, we developed the classical multigrid method on non-nested meshes. Then, we have pointed out the similarity between the Full MultiGrid (FMG) algorithm and the mesh adaptation algorithm. We have proposed a new Adaptive Full MultiGrid algorithm which improve the overall robustness of the adaptive process and its overall efficiency [25].

4.12. Metric-orthogonal and metric-aligned mesh adaptation

Participants: Frédéric Alauzet, Victorien Menier, Adrien Loseille [correspondant].

A new algorithm to derive adaptive meshes has been introduced through new cavity-based algorithms. It allows to generate anisotropic surface and volume mesh that are aligned along the eigenvector directions. This allows us to improv the quality of the meshes and to deal naturally with boundary layer mesh generation [19], [27].

GANG Project-Team

5. New Results

5.1. Highlights of the Year

Pierre Fraigniaud has received the Prize for Innovation in Distributed Computing 2014.

5.2. Graph and Combinatorial Algorithms

5.2.1. Collision-Free Network Exploration

In the collision-free exploration model considered in [16], a set of mobile agents is placed at different nodes of a n-node network. The agents synchronously move along the network edges in a collision-free way, i.e., in no round may two agents occupy the same node. In each round, an agent may choose to stay at its currently occupied node or to move to one of its neighbors. An agent has no knowledge of the number and initial positions of other agents. We are looking for the shortest possible time required to complete the collision-free *network exploration*, i.e., to reach a configuration in which each agent is guaranteed to have visited all network nodes and has returned to its starting location.

In this work, we first considered the scenario when each mobile agent knows the map of the network, as well as its own initial position. We established a connection between the number of rounds required for collision-free exploration and the degree of the minimum-degree spanning tree of the graph. We provided tight (up to a constant factor) lower and upper bounds on the collision-free exploration time in general graphs, and the exact value of this parameter for trees. For our second scenario, in which the network is unknown to the agents, we proposed collision-free exploration strategies running in $O(n^2)$ rounds for tree networks and in $O(n^5 \log n)$ rounds for general networks.

5.2.2. Properties of Graph Search Procedures

In [4], we study the last vertex discovered by a graph search such as BFS or DFS. End-vertices of a given graph search may have some nice properties (as for example it is well known that the last vertex of Lexicographic Breadth First Search (LBFS) in a chordal graph is simplicial). Therefore it is interesting to consider if these vertices can be recognized in polynomial time or not. A graph search is a mechanism for systematically visiting the vertices of a graph. At each step of a graph search, the key point is the choice of the next vertex to be explored. Graph searches only differ by this selection mechanism during which a tie-break rule is used. In this paper we study how the choice of the tie-break rule can determine the complexity of the end-vertex problem for BFS or DFS. In particular we prove a counter-intuitive NP-completeness result for Breadth First Search, answering a question of D.G. Corneil, E. Köhler and J-M Lanlignel.

5.2.3. Matchings in Hypergraphs

A rainbow matching for (not necessarily distinct) sets $F_1, ...F_k$ of hypergraph edges is a matching consisting of k edges, one from each F_i. The aim of [3] is twofold—to put order in the multitude of conjectures that relate to this concept (some first presented here), and to prove partial results on one of the central conjectures settled by Ryser, Brualdi and Stein.

5.2.4. Common Intervals and Application to Genome Comparison

In [6], we show how to identify generalized common and conserved nested intervals. This is a bio-informatics papers, explaining how to compute more relaxed variants of common or of conserved intervals of two permutations, which has applications in genome comparison. It also presents some properties of the family of intervals, useful for storing them.

5.2.5. Graph Decomposition

In [10], we present a general framework for computing a large family of graph decomposition, the H-join. It generalizes some well know tools like modular decomposition or split decomposition. The paper explains how to compute it in polynomial time. A new canonical decomposition for sesquiprime graphs is also presented.

5.2.6. Combinatorial Optimization

Normal cone and subdifferential have been generalized through various continuous functions; in [8], we focus on a non separable Q-subdifferential version. Necessary and sufficient optimality conditions for unconstrained nonconvex problems are revisited accordingly. For inequality constrained problems, Q-subdifferential and the lagrangian multipliers, enhanced as continuous functions instead of scalars, allow us to derive new necessary and sufficient optimality conditions. In the same way, the Legendre-Fenchel conjugate is generalized into Q-conjugate and global optimality conditions are derived by Q-conjugate as well, leading to a tighter inequality.

5.3. Distributed Computing

5.3.1. Rendezvous

5.3.1.1. Rendezvous of Anonymous Agents in Trees

In [5], we study the so-called *rendezvous problem* in the mobile agent setting in graph environments. In the studied model, two identical (anonymous) mobile agents start from arbitrary nodes of an unknown tree and have to meet at some node. Agents move in synchronous rounds: in each round an agent can either stay at the current node or move to one of its neighbors. We consider deterministic algorithms for this rendezvous task. The main result of our research is a tight trade-off between the optimal time of completing rendezvous and the size of memory of the agents. For agents with k memory bits, we show that optimal rendezvous time is $\Theta(n + n^2/k)$ in n-node trees. More precisely, if $k \geq c \log n$, for some constant c, we design agents accomplishing rendezvous in arbitrary trees of size n (unknown to the agents) in time $O(n + n^2/k)$, starting with arbitrary delay. We also show that no pair of agents can accomplish rendezvous in time $o(n + n^2/k)$, even in the class of lines of known length and even with simultaneous start. Finally, we prove that at least logarithmic memory is necessary for rendezvous, even for agents starting simultaneously in a n-node line.

5.3.1.2. Rendezvous of Distance-Aware Mobile Agents in Unknown Graphs

In [17], we study the problem of rendezvous of two mobile agents starting at distinct locations in an unknown graph. The agents have distinct labels and walk in synchronous steps. However, the graph is unlabeled and the agents have no means of marking the nodes of the graph and cannot communicate with or see each other until they meet at a node. When the graph is very large, we would like the time to rendezvous to be independent of the graph size and to depend only on the initial distance between the agents and some local parameters such as the degree of the vertices, and the size of the agent's label. It is well known that even for simple graphs of degree Δ, the rendezvous time can be exponential in Δ in the worst case. In this study, we introduce a new version of the rendezvous problem where the agents are equipped with a device that measures its distance to the other agent after every step. We show that these *distance-aware* agents are able to rendezvous in any unknown graph, in time polynomial in all the local parameters such the degree of the nodes, the initial distance D and the size of the smaller of the two agent labels $l = \min(l_1, l_2)$. Our algorithm has a time complexity of $O(\Delta(D + \log l))$ and we show an almost matching lower bound of $\Omega(\Delta(D + \log l / \log \Delta))$ on the time complexity of any rendezvous algorithm in our scenario. Further, this lower bound extends existing lower bounds for the general rendezvous problem without distance awareness.

5.3.1.3. Rendezvous of Heterogeneous Mobile Agents in Edge-Weighted Networks

In [22], we study the deterministic rendezvous problem in which a pair of heterogeneous agents, differing in the time required to traverse particular edges of the graph, need to meet on an edge or node of the graph. Each of the agents knows the complete topology of the undirected graph and the initial positions of both of the agents. The agent also knows its own traversal times for all of the edges of the graph, but is unaware of the corresponding traversal times for the other agent. In this scenario, we study the time required by the agents to meet, compared to the time T_{OPT} in the offline scenario in which the agents have complete knowledge of each others capabilities. When no additional assumptions are made, we show that rendezvous can be achieved after time $O(nT_{\mathrm{OPT}})$ in a n-node graph, and that this time is essentially the best possible in some cases. However, the rendezvous time can be reduced to $\Theta(T_{\mathrm{OPT}})$ when the agents are allowed to exchange $\Theta(n)$ bits of information at the start of the rendezvous process. We then show that under some natural assumption about the traversal times of edges, the hardness of the heterogeneous rendezvous problem can be substantially decreased, both in terms of time required for rendezvous without communication, and the communication complexity of achieving rendezvous in time $\Theta(T_{\mathrm{OPT}})$.

5.3.1.4. Rendezvous with Different Speeds

In [32] we introduce the study of the rendezvous problem in the context of agents having different speeds, and present tight and almost tight bounds for this problem, restricted to a ring topology.

5.3.2. Fair Synchronization

A non-blocking implementation of a concurrent object is an implementation that does not prevent concurrent accesses to the internal representation of the object, while guaranteeing the deadlock-freedom progress condition without using locks. Considering a failure free context, G. Taubenfeld has introduced (DISC 2013) a simple modular approach, captured under a new problem called the *fair synchronization* problem, to transform a non-blocking implementation into a starvation-free implementation satisfying a strong fairness requirement.

This approach is illustrated in [19] with the implementation of a concurrent stack. The spirit of the paper is mainly pedagogical. Its aim is not to introduce new concepts or algorithms, but to show that a powerful, simple, and modular transformation can provide concurrent objects with strong fairness properties.

In [20], we extend this approach in several directions. It first generalizes the fair synchronization problem to read/write asynchronous systems where any number of processes may crash. Then, it introduces a new failure detector and uses it to solve the fair synchronization problem when processes may crash. This failure detector, denoted QP (Quasi Perfect), is very close to, but strictly weaker than, the perfect failure detector. Last but not least, the paper shows that the proposed failure detector QP is optimal in the sense that the information on failures it provides to the processes can be extracted from any algorithm solving the fair synchronization problem in the presence of any number of process crash failures.

5.3.3. Wait Free with Advice

In [7], we motivate and propose a new way of thinking about failure detectors which allows us to define, quite surprisingly, what it means to solve a distributed task *wait-freeusing a failure detector*. In our model, the system is composed of *computation* processes that obtain inputs and are supposed to produce outputs and *synchronization* processes that are subject to failures and can query a failure detector.

Under the condition that *correct* synchronization processes take sufficiently many steps, they provide the computation processes with enough *advice* to solve the given task wait-free: every computation process outputs in a finite number of its own steps, regardless of the behavior of other computation processes.

Every task can thus be characterized by the *weakest* failure detector that allows for solving it, and we show that every such failure detector captures a form of set agreement. We then obtain a complete classification of tasks, including ones that evaded comprehensible characterization so far, such as renaming or weak symmetry breaking.

5.3.4. Adaptive Register Allocation

In [18], we give an adaptive algorithm in which processes use multi-writer multi-reader registers to acquire exclusive write access to their own single-writer, multi-reader registers. It is the first such algorithm that uses a number of registers linear in the number of participating processes. Previous adaptive algorithms require at least $\Theta(n^{3/2})$ registers.

5.3.5. Leader Election

Considering the case of homonyms processes (some processes may share the same identifier) on a ring [21], we give a necessary and sufficient condition on the number of identifiers to enable leader election. We prove that if l is the number of identifiers then message-terminating election is possible if and only if l is greater than the greatest proper divisor of the ring size even if the processes do not know the ring size. If the ring size is known, we propose a process-terminating algorithm exchanging $O(n\log(n))$ messages that is optimal.

5.3.6. Concurrency and Fault-tolerance

In [15], we study the connections between self-stabilization and proof-labeling schemes. It follows from the definition of *silent* self-stabilization, and from the definition of *proof-labeling* scheme, that if there exists a silent self-stabilizing algorithm using ℓ-bit registers for solving a task T, then there exists a proof-labeling scheme for T using registers of at most ℓ bits. The first result in this paper is the converse to this statement. We show that if there exists a proof-labeling scheme for a task T, using ℓ-bit registers, then there exists a silent self-stabilizing algorithm using registers of at most $O(\ell + \log n)$ bits for solving T, where n is the number of processes in the system. Therefore, as far as memory space is concerned, the design of silent self-stabilizing algorithms essentially boils down to the design of compact proof-labeling schemes. The second result in this paper addresses time complexity. We show that, for every task T with k-bits output size in n-node networks, there exists a silent self-stabilizing algorithm solving T in $O(n)$ rounds, using registers of $O(n^2 + kn)$ bits. Therefore, as far as running time is concerned, *every* task has a silent self-stabilizing algorithm converging in a linear number of rounds.

In [27], we study the connections between, on the one hand, asynchrony and concurrency, and, on the other hand, the quality of the expected solution of a distributed algorithm. The state machine approach is a well-known technique for building distributed services requiring high performance and high availability, by replicating servers, and by coordinating client interactions with server replicas using consensus. Indulgent consensus algorithms exist for realistic eventually partially synchronous models, that never violate safety and guarantee liveness once the system becomes synchronous. Unavoidably, these algorithms may never terminate, even when no processor crashes, if the system never becomes synchronous. We propose a mechanism similar to state machine replication, called *RC-simulation*, that can always make progress, even if the system is never synchronous. Using RC-simulation, the quality of the service will adjust to the current level of asynchrony of the network — degrading when the system is very asynchronous, and improving when the system becomes more synchronous. RC-simulation generalizes the state machine approach in the following sense: when the system is asynchronous, the system behaves as if $k+1$ threads were running concurrently, where k is a function of the asynchrony. In order to illustrate how the RC-simulation can be used, we describe a long-lived renaming implementation. By reducing the concurrency down to the asynchrony of the system, RC-simulation enables to obtain renaming quality that adapts linearly to the asynchrony.

5.3.7. Quantum Computing

In [1], we provide illustrative examples of distributed computing problems for which it is possible to design tight lower bounds for *quantum* algorithms without having to manipulate concepts from quantum mechanics, at all. As a case study, we address the following class of 2-player problems. Alice (resp., Bob) receives a boolean x (resp., y) as input, and must return a boolean a (resp., b) as output. A *game* between Alice and Bob is defined by a pair (δ, f) of boolean functions. The objective of Alice and Bob playing game (δ, f) is, for every pair (x, y) of inputs, to output values a and b, respectively, satisfying $\delta(a, b) = f(x, y)$, in *absence of any communication* between the two players, but in *presence of shared resources*. The ability of the two players to solve the game then depends on the type of resources they share. It is known that, for the so-called CHSH game, i.e., for the game $a \oplus b = x \wedge y$, the ability for the players to use entangled quantum bits (qubits) helps. We show that, apart from the CHSH game, quantum correlations do not help, in the sense that, for every game not equivalent to the CHSH game, there exists a classical protocol (using shared randomness) whose probability of success is at least as large as the one of any protocol using quantum resources. This result holds for both worst case and average case analysis. It is achieved by considering a model stronger than quantum correlations, the *non-signaling model*, which subsumes quantum mechanics, but is far easier to handle.

5.3.8. Distributed Decision and Verification

5.3.8.1. Randomization

In [12], we study the power of randomization in the context of locality by analyzing the ability to "boost" the success probability of deciding a distributed language. The main outcome of this analysis is that the distributed computing setting contrasts significantly with the sequential one as far as randomization is concerned. Indeed, we prove that in some cases, the ability to increase the success probability for deciding distributed languages is rather limited.

5.3.8.2. Model Variants

In a series of papers [14], [28], we analyze distributed decision in the context of various models for distributed computing.

In [28], we carry on the effort to bridging runtime verification with distributed computability, studying necessary conditions for monitoring failure prone asynchronous distributed systems. It has been recently proved that there are correctness properties that require a large number of opinions to be monitored, an opinion being of the form true, false, perhaps, probably true, probably no, etc. The main outcome of this paper is to show that this large number of opinions is not an artifact induced by the existence of artificial constructions. Instead, monitoring an important class of properties, requiring processes to produce at most k different values does require such a large number of opinions. Specifically, our main result is a proof that it is impossible to monitor k-set-agreement in an n-process system with fewer than $\min\{2k, n\} + 1$ opinions. We also provide an algorithm to monitor k-set-agreement with $\min\{2k, n\} + 1$ opinions, showing that the lower bound is tight.

Finally, in [14], we tackle *local distributed testing* of graph properties. This framework is well suited to contexts in which data dispersed among the nodes of a network can be collected by some central authority (like in, e.g., sensor networks). In local distributed testing, each node can provide the central authority with just a few information about what it perceives from its neighboring environment, and, based on the collected information, the central authority is aiming at deciding whether or not the network satisfies some property. We analyze in depth the prominent example of checking *cycle-freeness*, and establish tight bounds on the amount of information to be transferred by each node to the central authority for deciding cycle-freeness. In particular, we show that distributively testing cycle-freeness requires at least $\lceil \log d \rceil - 1$ bits of information per node in graphs with maximum degree d, even for connected graphs. Our proof is based on a novel version of the seminal result by Naor and Stockmeyer (1995) enabling to reduce the study of certain kinds of algorithms to order-invariant algorithms, and on an appropriate use of the known fact that every free group can be linearly ordered.

5.3.9. Voting Systems

In [44], [38], we consider a general framework for voting systems with arbitrary types of ballots such as orders of preference, grades, etc. We investigate their manipulability: in what states of the population may a coalition of electors, by casting an insincere ballot, secure a result that is better from their point of view?

We show that, for a large class of voting systems, a simple modification allows to reduce manipulability. This modification is *Condorcification*: when there is a Condorcet winner, designate her; otherwise, use the original rule.

When electors are independent, for any non-ordinal voting system (i.e. requiring information that is not included in the orders of preferences, for example grades), we prove that there exists an ordinal voting system whose manipulability rate is at most as high and which meets some other desirable properties. Furthermore, this result is also true when voters are not independent but the culture is *decomposable*, a weaker condition that we define.

Combining both results, we conclude that when searching for a voting system whose manipulability is minimal (in a large class of systems), one can restrict to voting systems that are ordinal and meet the Condorcet criterion.

In [35], we examine the geometrical properties of the space of expected utilities over a finite set of options, which is commonly used to model the preferences of an agent. We focus on the case where options are assumed to be symmetrical a priori, which is a classical neutrality assumption when studying voting systems. Specifically, we prove that the only Riemannian metric that respects the geometrical properties and the natural symmetries of the utility space is the round metric. Whereas Impartial Culture is widely used in Social Choice literature but limited to ordinal preference, our theoretical result allows to extend it canonically to cardinal preferences.

In [25], we study the manipulability of voting systems in a real-life experiment: electing the best paper in the conference Algotel 2012. Based on real ballots, we provide a quantitative study of the manipulability, as a function of the voting system used. We show that, even in a situation where all voting systems give the same winner by sincere voting, choosing the voting system is critical, because it has a huge impact on manipulability. In particular, one voting system fare way be better than the others: Instant-Runoff Voting.

5.4. Network Algorithms and Analysis

5.4.1. Bounds on the Cover Time in the Rotor-Router Model

In [23] and [33], we consider the *rotor-router mechanism*, which provides a deterministic alternative to the random walk in undirected graphs. In this model, a set of k identical walkers is deployed in parallel, starting from a chosen subset of nodes, and moving around the graph in synchronous steps. During the process, each node maintains a cyclic ordering of its outgoing arcs, and successively propagates walkers which visit it along its outgoing arcs in round-robin fashion, according to the fixed ordering. We consider the *cover time* of such a system, i.e., the number of steps after which each node has been visited by at least one walk, regardless of the starting locations of the walks. In the case of $k = 1$, Yanovski et al. (2003) and Bampas et al. (2009) showed that a single walk achieves a cover time of exactly $\Theta(mD)$ for any n-node graph with m edges and diameter D, and that the walker explores increasingly large Eulerian subgraphs before eventually stabilizes to a traversal of an Eulerian circuit on the set of all directed edges of the graph.

In [23], we provide tight bounds on the cover time of k parallel rotor walks in a graph. We show that this cover time is at most $\Theta(mD/\log k)$ and at least $\Theta(mD/k)$ for any graph, which corresponds to a speedup of between $\Theta(\log k)$ and $\Theta(k)$ with respect to the cover time of a single walk. Both of these extremal values of speedup are achieved for some graph classes. Our results hold for up to a polynomially large number of walks, $k = O(poly(n))$.

In [33], we perform a case study of cover time of the rotor-router, showing how the cover time depends on k for many important graph classes. We determine the precise asymptotic value of the rotor-router cover time for all values of k for degree-restricted expanders, random graphs, and constant-dimensional tori. For hypercubes, we also resolve the question precisely, except for values of k much larger than n. Our results can be compared to those obtained by Elsässer and Sauerwald (2009) in an analogous study of the cover time of k independent parallel random walks in a graph; for the rotor-router, we obtain tight bounds in a slightly broader spectrum of cases. Our proofs take advantage of a relation which we develop, linking the cover time of the rotor-router to the mixing time of the random walk and the local divergence of a discrete diffusion process on the considered graph.

5.4.2. Web Ranking and Aliveness

In [29] and [30], we investigate how to efficiently retrieve large portions of alive pages from an old crawl using orderings we called LiveRanks. Our work establishes the possibility of efficiently recovering a significant portion of the alive pages of an old snapshot and advocates for the use of an adaptive sample-based PageRank for obtaining an efficient LiveRank. Additionally, application field is not limited to Web graphs. It can be straightforwardly adapted to any online data with similar linkage enabling crawling, like P2P networks or online social networks.

5.4.3. Wireless Positioning

In [31], we consider how to construct a low-cost and efficient positioning system. We have proposed a new method called Two-Step Movement (2SM) to estimate the position of Mobile Terminal (MT). By exploiting useful information given by the position change of the device or user movement, this method can minimize the number of Reference Points (RP) required (*i.e.,* only one) in a localization system or navigation service and reduce system implementation cost. Analytical result shows that the user position can be derived, under noisy environment, with an estimation error about 10% of the distance between the RP and MT, or even less.

5.4.4. Content Centric Networking

Today's Internet usage is mostly centered around location-independent services. Because the Internet architecture is host-centric, content or service requests still have to be translated into locations, or the IP address of their hosts. This translation is realized through different technologies, e.g. DNS and HTTP redirection, which are currently implemented at the Application Layer. (ICN) proposes to evolve the current Internet infrastructure by extending the networking layer with name-based primitives.

In [45], we target the design and implementation of a content router, which is a network entity that implements *name-based forwarding*, or it can forward packets based on the content name they are addressed to. This work makes three major contributions. First, we propose an algorithm for name-based longest prefix match whose main novelty is the *prefix Bloom filter*, a Bloom filter variant that exploits the hierarchical nature of content prefixes. Second, a content router design that is compatible with both today's networking protocols and with widely used network equipments. Third, two innovative features that increase the scalability of a content router both in term of forwarding-information-base size and forwarding speed.

In the demonstration [34] held in the ICN conference, we demonstrate a high speed Information-Centric Network in a mobile backhaul setting. In particular, we emulate an information aware data plane and we highlight the significant benefits it provides in terms of both user experience and network provider cost in the backhaul setting. Our setup consists of high-speed ICN devices employed in a down-scaled realistic representation of a mobile backhaul topology, fed with traffic workloads characterized from Orange's mobile network. We compare numerical results activating and de-activating the ICN feature at run-time, showing the main differences between the two approaches. All the devices are implemented in a real high-speed multi-core equipment, and they are connected by means of internal port connections. Traffic is injected using a Traffic Generator which is implemented in the same architecture.

5.4.5. Information Dissemination

5.4.5.1. Dissemination with Noise or Limited Memory

In [26], we introduce the study of basic distributed computing problems in the context of noise in communication. We establish tight and almost tight bounds for the rumor spreading problem as well as for the majority-consensus problem.

In [11], we theoretically study a general model of information sharing within animal groups. We take an algorithmic perspective to identify efficient communication schemes that are, nevertheless, economic in terms of communication, memory and individual internal computation. We present a simple and natural algorithm in which each agent compresses all information it has gathered into a single parameter that represents its confidence in its behavior. Confidence is communicated between agents by means of active signaling. We motivate this model by novel and existing empirical evidences for confidence sharing in animal groups. We rigorously show that this algorithm competes extremely well with the best possible algorithm that operates without any computational constraints. We also show that this algorithm is minimal, in the sense that further reduction in communication may significantly reduce performances. Our proofs rely on the Cramér-Rao bound and on our definition of a Fisher Channel Capacity. We use these concepts to quantify information flows within the group which are then used to obtain lower bounds on collective performance.

5.4.5.2. Gossip and Rumor Spreading with Flooding

In [2], we address the flooding problem in dynamic graphs, where flooding is the basic mechanism in which every node becoming aware of an information at step t forwards this information to all its neighbors at all forthcoming steps $t' > t$. In particular, we show that a technique developed in a previous paper, for analyzing flooding in a Markovian sequence of Erdös-Rényi graphs, is robust enough to be used also in different contexts. We establish this by analyzing flooding in a sequence of graphs drawn independently at random according to a model of random graphs with given expected degree sequence. In the prominent case of power-law degree distributions, we prove that flooding takes almost surely $O(\log n)$ steps even if, almost surely, none of the graphs in the sequence is connected. In the general case of graphs with an arbitrary degree sequence, we prove several upper bounds on the flooding time, which depend on specific properties of the degree sequence.

5.4.6. Small-world Networks

In [9], we study decentralized routing in small-world networks that combine a wide variation in node degrees with a notion of spatial embedding. Specifically, we consider a variant of J. Kleinberg's grid-based small-world model in which (1) the number of long-range edges of each node is not fixed, but is drawn from a power-law probability distribution with exponent parameter $\alpha \geq 0$ and constant mean, and (2) the long-range edges are considered to be bidirectional for the purposes of routing. This model is motivated by empirical observations indicating that several real networks have degrees that follow a power-law distribution. The measured power-law exponent α for these networks is often in the range between 2 and 3. For the small-world model we consider, we show that when $2 < \alpha < 3$ the standard greedy routing algorithm, in which a node forwards the message to its neighbor that is closest to the target in the grid, finishes in an expected number of $O(\log^{\alpha-1} n \cdot \log\log n)$ steps, for any source–target pair. This is asymptotically smaller than the $O(\log^2 n)$ steps needed in Kleinberg's original model with the same average degree, and approaches $O(\log n)$ as α approaches 2. Further, we show that when $0 \leq \alpha < 2$ or $\alpha \geq 3$ the expected number of steps is $O(\log^2 n)$, while for $\alpha = 2$ it is $O(\log^{4/3} n)$. We complement these results with lower bounds that match the upper bounds within at most a $\log\log n$ factor.

5.4.7. Voting Systems and Path Selection in Networks

In [24], we apply our theoretical and experimental results on voting systems to a network use case: choosing a path in a network. In our model, nodes have an economical reward or cost for each possible path and they vote to elect the path. We show that the choice of the voting system has an important impact on the manipulability and the economical efficiency of this system. From both points of view, Instant-Runoff Voting gives the best results.

GCG Team

6. New Results

6.1. Highlights of the Year

Graduate Research Award of the OSU department in 2015 for Venmugil Elango (co-advised by Fabrice Rastello)

6.2. An interval constrained memory allocator for a GAS runtime

Participants: François Gindraud, Fabrice Rastello, Albert Cohen [ENS Ulm].

This work presents a memory allocator for global address space (GAS) runtime targeting distributed memory embedded architectures (MPSoC). MPSoC we are interested in are relatively new architectures, composed of several nodes with multiple general purpose cores and a local memory, linked by a network, all on one chip (NoC). They have promising energy and computing performances, but are hard to program due to the multilevel parallelism and the hardware constraints (limited memory, network structure). Existing programming framework are either thin but let the programmer do the hard choices (OpenMP + MPI) or heavy and automatic but target specific kind of applications on big systems (Global Arrays).

Givy 5.1 is a runtime currently developed to execute dynamic task graphs with data-flow dependencies on MPSoC. It has a focus on supporting irregular applications, using the dependencies to perform data-aware dynamic task scheduling and data transfer. Data blocks live in a GAS, and thus requires a GAS-aware memory allocator to avoid address collisions when they are dynamically allocated. The allocator implementation proposed in this paper does this with zero synchronization between nodes, while being memory efficient in the small distributed memories, and fast on each multithreaded node.

This work will be submitted at ACM ISMM Symposium.

6.3. A Framework for Enhancing Data Reuse via Associative Reordering

Participants: Kevin Stock [OSU], Martin Kong [OSU], Tobias Grosser [ENS Ulm], Louis-Noël Pouchet [UCLA], Fabrice Rastello, J. Ramanujam [LSU], P. Sadayappan [OSU].

The freedom to reorder computations involving associative operators has been widely recognized and exploited in designing parallel algorithms and to a more limited extent in optimizing compilers.

In this work, we develop a novel framework utilizing the associativity and commutativity of operations in regular loop computations to enhance register reuse. Stencils represent a particular class of important computations where the optimization framework can be applied to enhance performance. We show how stencil operations can be implemented to better exploit register reuse and reduce load/stores. We develop a multi-dimensional retiming formalism to characterize the space of valid implementations in conjunction with other program transformations. Experimental results demonstrate the effectiveness of the framework on a collection of high-order stencils.

This work is the fruit of the collaboration 8.1 with OSU and has been presented at the conference ACM PLDI'14.

6.4. Beyond Reuse Distance Analysis: Dynamic Analysis for Characterization of Data Locality Potential

Participants: Naznin Fauzia [OSU], Venmugil Elango [OSU], Mahesh Ravishankar [OSU], J. Ramanujam [LSU], Fabrice Rastello, Atanas Routnev [OSU], Louis-Noël Pouchet [UCLA], P. Sadayappan [OSU].

Emerging computer architectures will feature drastically decreased flops/byte (ratio of peak processing rate to memory bandwidth) as highlighted by recent studies on Exascale architectural trends. Further, flops are getting cheaper while the energy cost of data movement is increasingly dominant. The understanding and characterization of data locality properties of computations is critical in order to guide efforts to enhance data locality.

Reuse distance analysis of memory address traces is a valuable tool to perform data locality characterization of programs. A single reuse distance analysis can be used to estimate the number of cache misses in a fully associative LRU cache of any size, thereby providing estimates on the minimum bandwidth requirements at different levels of the memory hierarchy to avoid being bandwidth bound. However, such an analysis only holds for the particular execution order that produced the trace. It cannot estimate potential improvement in data locality through dependence preserving transformations that change the execution schedule of the operations in the computation.

In this work, we develop a novel dynamic analysis approach to characterize the inherent locality properties of a computation and thereby assess the potential for data locality enhancement via dependence preserving transformations.

This work is the fruit of the collaboration 8.1 with OSU and has been published at ACM TACO'14.

6.5. On Using the Roofline Model with Lower Bounds on Data Movement

Participants: Venmugil Elango [OSU], Naser Sedaghati [OSU], Fabrice Rastello, Louis-Noël Pouchet [UCLA], J. Ramanujam [LSU], Radu Teodorescu [OSU], P. Sadayappan [OSU].

The roofline model is a popular approach to "bounds and bottleneck" performance analysis. It focuses on the limits to performance of processors because of limited bandwidth to off-chip memory. It models upper bounds on performance as a function of operational intensity, the ratio of computational operations per byte of data moved from/to memory. While operational intensity can be directly measured for a specific implementation of an algorithm on a particular target platform, it is of interest to obtain broader insights on bottlenecks, where various semantically equivalent implementations of an algorithm are considered, along with analysis for variations in architectural parameters. This is currently very cumbersome and requires performance modeling and analysis of many variants.

In this work, we alleviate this problem by using the roofline model in conjunction with upper bounds on the operational intensity of computations as a function of cache capacity, derived using lower bounds on data movement. This enables bottleneck analysis that holds across all dependence-preserving semantically equivalent implementations of an algorithm. We demonstrate the utility of the approach in in assessing fundamental limits to performance and energy efficiency for several benchmark algorithms across a design space of architectural variations.

This work is the fruit of the collaboration 8.1 with OSU and is to be published at ACM TACO'15.

6.6. On Characterizing the Data Access Complexity of Programs

Participants: Venmugil Elango [OSU], Fabrice Rastello, Louis-Noël Pouchet [UCLA], J. Ramanujam [LSU], P. Sadayappan [OSU].

Technology trends will cause data movement to account for the majority of energy expenditure and execution time on emerging computers. Therefore, computational complexity will no longer be a sufficient metric for comparing algorithms, and a fundamental characterization of data access complexity will be increasingly important. The problem of developing lower bounds for data access complexity has been modeled using the formalism of Hong & Kung's red/blue pebble game for computational directed acyclic graphs (CDAGs). However, previously developed approaches to lower bounds analysis for the red/blue pebble game are very limited in effectiveness when applied to CDAGs of real programs, with computations comprised of multiple sub-computations with differing DAG structure. We address this problem by developing an approach for effectively composing lower bounds based on graph decomposition. We also develop a static analysis algorithm to derive the asymptotic data-access lower bounds of programs, as a function of the problem size and cache size.

This work is the fruit of the collaboration 8.1 with OSU and is to be presented at ACM POPL'15.

6.7. PolyCheck: Dynamic Verification of Iteration Space Transformations on Affine Programs

Participants: Sriram Krishnamoorthy [PNNL], Bao Wenlei [OSU], Louis-Noël Pouchet [UCLA], P. Sadayappan [OSU], Fabrice Rastello.

High-level compiler transformations, especially loop transformations, are widely recognized as critical optimizations to restructure programs to improve data locality and expose parallelism.

Guaranteeing the correctness of program transformations is essential, and to date three main approaches have been developed: proof of equivalence of affine programs, matching the execution traces of programs, and checking bit-by-bit equivalence of the outputs of the programs. Each technique suffers from limitations in either the kind of transformations supported, space complexity, or the sensitivity to the testing dataset. In this paper, we take a novel approach addressing all three limitations to provide an automatic bug checker to verify any iteration reordering transformations on affine programs, including non-affine transformations, with space consumption proportional to the original program data, and robust to arbitrary datasets of a given size. We achieve this by exploiting the structure of affine program control- and data-flow to generate at compile-time a lightweight checker code to be executed within the transformed program. Experimental results assess the correctness and effectiveness of our method, and its increased coverage over previous approaches.

This work is the result of the collaboration 8.1 with OSU.

6.8. On Using Lower Bounds for Discrimination of Utility/Futility of Loop Fusion

Participants: Samyam Rajbhandari [OSU], Martin Konk [OSU], P. Sadayappan [OSU], Robert J. Harrison [Stonybrook], Fabrice Rastello.

Fusion is an important loop transformation for data locality enhancement. However, it is very challenging to determine which of a set of possible fusion choices is best. In this paper, we pursue a novel approach to addressing this problem. Instead of the conventional approach of explicitly modeling different possible fused loop configurations and modeling the expected performance with each, we instead use lower bounds modeling to characterize conditions where fusion might have utility and where it will be futile because the maximal possible improvement from fusion is much lower than the minimal data movement overheads for each of the unfused components. We successfully demonstrate the use of such a methodology with two practically important codes from the quantum chemistry domain, i) with the affine 4-index transform code, and ii) unstructured tree operations with the MADNESS framework.

This work is the result of the collaboration 8.1 with OSU.

6.9. A Tiling Perspective for Register Optimization

Participants: Duco Van Amstel, Lukasz Domagala, P. Sadayappan [OSU], Fabrice Rastello.

Register allocation is a much studied problem. A particularly important context for optimizing register allocation is within loops, since a significant fraction of the execution time of programs is often inside loop code. A variety of algorithms have been proposed in the past for register allocation, but the complexity of the problem has resulted in a decoupling of several important aspects, including loop unrolling, register promotion, and instruction reordering.

In this work, we develop an approach to register allocation and promotion in a unified optimization framework that simultaneously considers the impact of loop unrolling and instruction scheduling. This is done via a novel instruction tiling approach where instructions within a loop are represented along one dimension and innermost loop iterations along the other dimension. By exploiting the regularity along the loop dimension, and imposing essential dependence based constraints on intra-tile execution order, the problem of optimizing register pressure is cast in a constraint programming formalism. Experimental results are provided from thousands of innermost loops extracted from the SPEC benchmarks, demonstrating improvements over the current state-of-the-art.

This work is the fruit of both the collaboration 8.1 with OSU and with Kalray 7.1 7.2 .

6.10. Hybrid Pointer Disambiguation

Participants: Fernando Pereira, Alexandros Labrineas, Péricles Alves, Fabian Gruber, Fabrice Rastello.

In order to provide effective optimizations, compilers must deal with memory dependences. However, the state-of-the-art heuristics available in the literature to track memory dependencies are inherently imprecise and computationally expensive. Consequently, the most advanced code transformations that compilers have today are ineffective when applied on real-world programs. The goal of this paper is to solve this conundrum - a goal that we accomplish through the hybrid disambiguation of pointers. We provide a static analysis that generates dynamic tests to determine when two memory locations can overlap. We then produce two versions of a loop: one that is aliasing-free - hence, easy to optimize - and another that is not. Our checks lets us safely branch to the optimizable region. We have applied these ideas on Polly-LLVM, a loop optimizer built on top of the LLVM compilation infrastructure. Our experiments indicate that our method is precise, effective and useful: we can disambiguate the vast majority of checks in benchmarks that go from the loop intensive Polybench suite to the more general SPEC CPU 2006 benchmark collection. The result of this precision is code quality: the binaries that we generate are 9.5% faster than those that Polly-LLVM produces without our optimization. Given the current technology to statically solve alias analysis, we believe that our ideas are a necessary step to make modern compiler optimizations useful in practice.

This work is the fruit of the collaboration with UFMG 8.1 and Kalray 7.1 7.2 .

6.11. Parameterized Construction of Program Representations for Sparse Dataflow Analyses

Participants: André Tavares [ENS Lyon], Benoit Boissinot [ENS Lyon], Fernando Pereira, Fabrice Rastello.

Data-flow analyses usually associate information with control flow regions. Informally, if these regions are too small, like a point between two consecutive statements, we call the analysis dense. On the other hand, if these regions include many such points, then we call it sparse. This paper presents a systematic method to build program representations that support sparse analyses. To pave the way to this framework we clarify the bibliography about well-known intermediate program representations. We show that our approach, up to parameter choice, subsumes many of these representations, such as the SSA, SSI and e-SSA forms. In particular, our algorithms are faster, simpler and more frugal than the previous techniques used to construct SSI - Static Single Information - form programs. We produce intermediate representations isomorphic to Choi *et al.*'s Sparse Evaluation Graphs (SEG) for the family of data-flow problems that can be partitioned per variables. However, contrary to SEGs, we can handle - sparsely - problems that are not in this family. We have tested our ideas in the LLVM compiler, comparing different program representations in terms of size and construction time.

This work is the fruit of the collaboration with UFMG 8.1 and has been presented at Springer CC'14.

6.12. Time-critical Computing on a Single-chip Massively Parallel Processor

Participants: Benoit Dupont-de-Dinechin [Kalray], Duco Van Amstel, Marc Poulhiès [Kalray], Guillaume Lager [Kalray].

In this work we demonstrate the capabilities of the MPPA(TM)-256 chip in the field of time-critical computations. This manycore chip features amongst others a Network-on-Chip (NoC) linking the seperate computational clusters each disposing of its own local memory and processing elements (PEs). The PEs architectural features induce a locally deterministic behaviour and the memory access arbitration that is used allows for a Worst-Case Execution Time (WCET) that is achieved for the combination of all local worst-cases. As such, in order to achieve a WCET analysis for a full MPPA(TM)-256 chip, we provide a Worst-Case Traversal Time (WCTT) analysis for the NoC to link the WCETs provided by each computational cluster. This part of the work is based on the (sigma, rho) model used for general network flow analysis and Quality-of-Service (QoS) parametrization.

This work has been presented at DATE'14.

6.13. Guaranteed Services of the NoC of a Manycore Processor

Participants: Benoit Dupont-de-Dinechin [Kalray], Yves Durand [CEA], Duco Van Amstel [Kalray], Alexandre Ghiti [Kalray].

In the case of the MPPA(TM)-256 chip the study of the integrated Network-on-Chip (NoC) is a fundamental subject for anyone using this architecture for time-critical purposes or real-time use-cases that need guarantees on the Worst-Case Traversal Time (WCTT) of the NoC. Previous work has already shown that the MPPA(TM)-256 NoC can be modelled using the (sigma, rho)-model. In the current work we will elaborate on this point by providing an indepth analysis of the NoC as well as the method to guarantee Quality-of-Service properties.

This work has been presented at the International Workshop on Network on Chip Architectures 2014.

GECO Project-Team

6. New Results

6.1. Highlights of the Year

We organized a thematic trimester on “Geometry, analysis and dynamics on sub-Riemannian manifolds” at the Institut Henri Poincaré (IHP), including 4 workshops, 4 research courses, 8 thematic days, several seminars. We also organized an associated school at CIRM with 4 introductory courses. The web pages of the events are:
http://www.cmap.polytechnique.fr/subriemannian/
http://www.cmap.polytechnique.fr/subriemannian/cirm/

6.2. New results: geometric control

Let us list some new results in sub-Riemannian geometry and hypoelliptic diffusion obtained by GECO’s members.

- The article [14] presents simple controls that generate motion in the direction of high order Lie brackets. Whereas the naive use of piecewise constant controls requires the number of switchings to grow exponentially with the length of the bracket, we show that such motion is possible with sinusoidal controls whose sum of frequencies equals the length of the bracket. This work is closely related and motivated by the study of the complexity of sub-Riemannian geodesics for generic regular distributions, i.e., whose derived flag has maximal growth vector. Of particular interest is the approximation of curves transversal to the distribution by admissible curves. We also present a surprising example that shows that it is possible to simultaneously kill higher moments without increasing the number of self-intersections of the base curve.
- The curvature discussed in [18] is a rather far going generalization of the Riemann sectional curvature. We define it for a wide class of optimal control problems: a unified framework including geometric structures such as Riemannian, sub-Riemannian, Finsler and sub-Finsler structures; a special attention is paid to the sub-Riemannian (or Carnot-Caratheodory) metric spaces. Our construction of the curvature is direct and naive, and it is similar to the original approach by Riemann. Surprisingly, it works in a very general setting and, in particular, for all sub-Riemannian spaces.
- In [19] we prove sectional and Ricci-type comparison theorems for the existence of conjugate points along sub-Riemannian geodesics. In order to do that, we regard sub-Riemannian structures as a special kind of variational problems. In this setting, we identify a class of models, namely linear quadratic optimal control systems, that play the role of the constant curvature spaces. As an application, we prove a version of sub-Riemannian Bonnet–Myers theorem and we obtain some new results on conjugate points for 3D left-invariant sub-Riemannian structures.
- In the study of conjugate times in sub-Riemannian geometry, linear quadratic optimal control problems show up as model cases. In [1] we consider a dynamical system with a constant, quadratic Hamiltonian h, and we characterize the number of conjugate times in terms of the spectrum of the Hamiltonian vector field H. We prove the following dichotomy: the number of conjugate times is identically zero or grows to infinity. The latter case occurs if and only if H has at least one Jordan block of odd dimension corresponding to a purely imaginary eigenvalue. As a byproduct, we obtain bounds from below on the number of conjugate times contained in an interval in terms of the spectrum of H.

- A 3D almost-Riemannian manifold is a generalized Riemannian manifold defined locally by 3 vector fields that play the role of an orthonormal frame, but could become collinear on some set called the singular set. Under the Hormander condition, a 3D almost-Riemannian structure still has a metric space structure, whose topology is compatible with the original topology of the manifold. Almost-Riemannian manifolds were deeply studied in dimension 2. In [21] we start the study of the 3D case which appear to be reacher with respect to the 2D case, due to the presence of abnormal extremals which define a field of directions on the singular set. We study the type of singularities of the metric that could appear generically, we construct local normal forms and we study abnormal extremals. We then study the nilpotent approximation and the structure of the corresponding small spheres. We finally give some preliminary results about heat diffusion on such manifolds.
- In [22] we study spectral properties of the Laplace-Beltrami operator on two relevant almost-Riemannian manifolds, namely the Grushin structures on the cylinder and on the sphere. As for general almost-Riemannian structures (under certain technical hypothesis), the singular set acts as a barrier for the evolution of the heat and of a quantum particle, although geodesics can cross it. This is a consequence of the self-adjointness of the Laplace-Beltrami operator on each connected component of the manifolds without the singular set. We get explicit descriptions of the spectrum, of the eigenfunctions and their properties. In particular in both cases we get a Weyl law with dominant term $E \log E$. We then study the effect of an Aharonov-Bohm non-apophantic magnetic potential that has a drastic effect on the spectral properties. Other generalized Riemannian structures including conic and anti-conic type manifolds are also studied. In this case, the Aharonov-Bohm magnetic potential may affect the self-adjointness of the Laplace-Beltrami operator.
- In [28] we investigate the number of geodesics between two points p and q on a contact sub-Riemannian manifold M. We show that the count of geodesics on M is controlled by the count on its nilpotent approximation at p (a contact Carnot group). For contact Carnot groups we give sharp bounds for a generic point q. Removing the genericity condition for q, geodesics might appear in families and we prove a similar statement for their topology. We study these families, and in particular we focus on the unexpected appearance of isometrically non-equivalent geodesics: families on which the action of isometries is not transitive. We apply the previous study to contact sub-Riemannian manifolds: we prove that for any given point $p \in M$ there is a sequence of points p_n such that $p_n \to p$ and that the number of geodesics between p and p_n grows unbounded (moreover these geodesics have the property of being contained in a small neighborhood of p).

New results on automatic control and motion planning for various type of applicative domains are the following.

- [8] is devoted to the problem of model-based prognostics for a Waste Water Treatment Plant (WWTP). Our aim is to predict degradation of certain parameters in the process, in order to anticipate malfunctions and to schedule maintenance. It turns out that a WWTP, together with the possible malfunction, has a specific structure: mostly, the malfunction appears in the model as an unknown input function. The process is observable whatever this unknown input is, and the unknown input can itself be identified through the observations. Due to this property, our method does not require any assumption of the type “slow dynamics degradation", as is usually assumed in ordinary prognostic methods. Our system being unknown-input observable, standard observer-based methods are enough to solve prognostic problems. Simulation results are shown for a typical WWTP.
- In [9] we study the problem of controlling an unmanned aerial vehicle (UAV) to provide a target supervision and/or to provide convoy protection to ground vehicles. We first present a control strategy based upon a Lyapunov-LaSalle stabilization method to provide supervision of a stationary target. The UAV is expected to join a predesigned admissible circular trajectory around the target which is itself a fixed point in the space. Our strategy is presented for both high altitude long endurance (HALE) and medium altitude long endurance (MALE) types of UAVs.
- In [12] we study how a particular spatial structure with a buffer impacts the number of equilibria and their stability in the chemostat model. We show that the occurrence of a buffer can allow a

species to setup or on the opposite to go to extinction, depending on the characteristics of the buffer. For non-monotonic response function, we characterize the buffered configurations that make the chemostat dynamics globally asymptotically stable, while this is not possible with single, serial or parallel vessels of the same total volume and input flow. These results are illustrated with the Haldane kinetic function.

- In [15] and [25] we present new results on the path planning problem in the case study of the car with trailers. We formulate the problem in the framework of optimal nonholonomic interpolation and we use standard techniques of nonlinear optimal control theory for deriving hyperelliptic signals as controls for driving the system in an optimal way. The hyperelliptic curves contain as many loops as the number of nonzero Lie brackets generated by the system. We compare the hyperelliptic signals with the ordinary Lissajous-like signals that appear in the literature, we conclude that the former have better performance.
- In [27] we consider affine-control systems, i.e., systems in the form $\dot{q}(t) = f_0(q(t)) + \sum_{i=1}^{m} u_i(t) f_i(q(t))$. Here, the point q belongs to a smooth manifold M, the f_i's are smooth vector fields on M. This type of system appears in many applications for mechanical systems, quantum control, microswimmers, neuro-geometry of vision...

We conclude the section by mentioning the book [17] that we edited, collecting some papers in honour of Andrei A. Agrachev for his 60th birthday. The book contains new results on sub-Riemannian geometry and more generally on the geometric theory of control.

6.3. New results: quantum control

New results have been obtained for the control of the bilinear Schrödinger equation.

- In [2] we present a sufficient condition for approximate controllability of the bilinear discrete-spectrum Schrödinger equation in the multi-input case. The controllability result extends to simultaneous controllability, approximate controllability in H^s, and tracking in modulus. The sufficient condition is more general than those present in the literature even in the single-input case and allows the spectrum of the uncontrolled operator to be very degenerate (e.g. to have multiple eigenvalues or equal gaps among different pairs of eigenvalues). We apply the general result to a rotating polar linear molecule, driven by three orthogonal external fields. A remarkable property of this model is the presence of infinitely many degeneracies and resonances in the spectrum.
- In [5] we consider the minimum time population transfer problem for a two level quantum system driven by two external fields with bounded amplitude. The controls are modeled as real functions and we do not use the Rotating Wave Approximation. After projection on the Bloch sphere, we treat the time-optimal control problem with techniques of optimal synthesis on 2D manifolds. Based on the Pontryagin Maximum Principle, we characterize a restricted set of candidate optimal trajectories. Properties on this set, crucial for complete optimal synthesis, are illustrated by numerical simulations. Furthermore, when the two controls have the same bound and this bound is small with respect to the difference of the two energy levels, we get a complete optimal synthesis up to a small neighborhood of the antipodal point of the initial condition.
- In [11] we investigate the controllability of quantum electrons trapped in a two-dimensional device, typically a metal oxide semiconductor (MOS) field-effect transistor. The problem is modeled by the Schrödinger equation in a bounded domain coupled to the Poisson equation for the electrical potential. The controller acts on the system through the boundary condition on the potential, on a part of the boundary modeling the gate. We prove that, generically with respect to the shape of the domain and boundary conditions on the gate, the device is controllable. We also consider control properties of a more realistic nonlinear version of the device, taking into account the self-consistent electrostatic Poisson potential.
- In [29] we prove the approximate controllability of a bilinear Schrödinger equation modelling a two trapped ions system. A new spectral decoupling technique is introduced, which allows to analyze the controllability of the infinite-dimensional system through finite-dimensional considerations.

6.4. New results: neurophysiology

- [3] presents a semidiscrete alternative to the theory of neurogeometry of vision, due to Citti, Petitot, and Sarti. We propose a new ingredient, namely, working on the group of translations and discrete rotations $SE(2,N)$. The theoretical side of our study relates the stochastic nature of the problem with the Moore group structure of $SE(2,N)$. Harmonic analysis over this group leads to very simple finite dimensional reductions. We then apply these ideas to the inpainting problem which is reduced to the integration of a completely parallelizable finite set of Mathieu-type diffusions (indexed by the dual of $SE(2,N)$ in place of the points of the Fourier plane, which is a drastic reduction). The integration of the the Mathieu equations can be performed by standard numerical methods for elliptic diffusions and leads to a very simple and efficient class of inpainting algorithms. We illustrate the performances of the method on a series of deeply corrupted images.
- In [4] and [7] we consider the problem of minimizing $\int_0^l \sqrt{\xi^2 + K(s)^2} ds$ for a planar curve having fixed initial and final positions and directions. The total length l is free. Here s is the arclength parameter, $K(s)$ is the curvature of the curve and $\xi > 0$ is a fixed constant. This problem comes from a model of geometry of vision due to Petitot, Citti and Sarti. We study existence of local and global minimizers for this problem. In [7] we characterize sub-Riemannian geodesics and the range of the exponential map. In [4] we prove that if for a certain choice of boundary conditions there is no global minimizer, then there is neither a local minimizer nor a geodesic. We finally give properties of the set of boundary conditions for which there exists a solution to the problem.

6.5. New results: switched systems

- In [6] we consider a family of linear control systems $\dot{x} = Ax + \alpha Bu$ on $\mathbb{R}^d$, where α belongs to a given class of persistently exciting signals. We seek maximal α-uniform stabilization and destabilization by means of linear feedbacks $u = Kx$. We extend previous results obtained for bidimensional single-input linear control systems to the general case as follows: if there exists at least one K such that the Lie algebra generated by A and BK is equal to the set of all $d \times d$ matrices, then the maximal rate of convergence of (A, B) is equal to the maximal rate of divergence of (A, B). We also provide more precise results in the general single-input case, where the above result is obtained under the simpler assumption of controllability of the pair (A, B).
- The paper [10] considers the stabilization to the origin of a persistently excited linear system by means of a linear state feedback, where we suppose that the feedback law is not applied instantaneously, but after a certain positive delay (not necessarily constant). The main result is that, under certain spectral hypotheses on the linear system, stabilization by means of a linear delayed feedback is indeed possible, generalizing a previous result already known for non-delayed feedback laws.
- In [16] and [26] we give a collection of converse Lyapunov–Krasovskii theorems for uncertain retarded differential equations. We show that the existence of a weakly degenerate Lyapunov–Krasovskii functional is a necessary and sufficient condition for the global exponential stability of the linear retarded functional differential equations. This is carried out using the switched system transformation approach.
- Consider a continuous-time linear switched system on $\mathbb{R}^n$ associated with a compact convex set of matrices. When it is irreducible and its largest Lyapunov exponent is zero there always exists a Barabanov norm associated with the system. In [23] we deal with two types of issues: (a) properties of Barabanov norms such as uniqueness up to homogeneity and strict convexity; (b) asymptotic behaviour of the extremal solutions of the linear switched system. Regarding Issue (a), we provide partial answers and propose four related open problems. As for Issue (b), we establish, when $n = 3$, a Poincaré-Bendixson theorem under a regularity assumption on the set of matrices. We then revisit a noteworthy result of N.E. Barabanov describing the asymptotic behaviour of linear switched system on $\mathbb{R}^3$ associated with a pair of Hurwitz matrices $\{A, A + bc^T\}$. After pointing out a fatal gap in Barabanov's proof we partially recover his result by alternative arguments.

- In [24] we address the exponential stability of a system of transport equations with intermittent damping on a network of $N \geq 2$ circles intersecting at a single point O. The N equations are coupled through a linear mixing of their values at O, described by a matrix M. The activity of the intermittent damping is determined by persistently exciting signals, all belonging to a fixed class. The main result is that, under suitable hypotheses on M and on the rationality of the ratios between the lengths of the circles, such a system is exponentially stable, uniformly with respect to the persistently exciting signals. The proof relies on an explicit formula for the solutions of this system, which allows one to track down the effects of the intermittent damping.

GENSCALE Project-Team

6. New Results

6.1. Highlights of the Year

discoSnp published in NAR. The publication presents a wide range of discoSnp applications that highlight the advantages and the drawbacks of predicting SNPs when no reference genomes are available. The publication witnesses the enthusiasm of users regarding both reference-free methods and the quality of the method. [20]

6.2. NGS methodology

Participants: Erwan Drezen, Anaïs Gouin, Dominique Lavenier, Claire Lemaitre, Antoine Limasset, Pierre Peterlongo, Guillaume Rizk.

Comparison of large sets of metagenomics data
We significantly extend the previous method (implemented in the Comparead tool) for computing similarity between sets of raw non assembled (and usually non-assemblable with current state of the art assemblers) reads. This enhancement of the method enables computations to be factorized when N read sets have to be compared all together. Moreover, the great advantage of this improvment is to save huge disk space and to enable efficient logical operations between metagenomic subset of reads. The Commet tool implements this optimized version.[25]

De novo SNP discovery
We developed a very efficient new way for detecting isolated SNPs given one, two or more raw read set(s) without using any reference genome. The implementation, called discoSnp, was applied to various datasets and applications. In particular, compared to finding isolated SNPs using a state-of-the-art assembly and mapping approach, our method requires significantly less computational resources, shows similar precision/recall values, and highly ranked predictions are less likely to be false positives. An experimental validation was conducted on an arthropod species (the tick *Ixodes ricinus*) on which de novo sequencing was performed. Among the predicted SNPs that were tested, 96% were successfully genotyped and truly exhibited polymorphism. [20]

De novo discovery of inversion breakpoints
A formal model has been proposed, together with an algorithm, for detecting inversion breakpoints without a reference genome, directly from raw NGS data. This model is characterized by a fixed size topological pattern in the de Bruijn Graph. We describe precisely the possible sources of false positives and false negatives and we additionally propose a sequence-based filter giving a good trade-off between precision and recall of the method. We implemented these ideas in a software called TakeABreak. Applied on simulated inversions in genomes of various complexity (from E. coli to a human chromosome dataset), the method provided promising results with a low memory footprint and a small computational time. [24]

Integrated detection and assembly of long insertion variants
We investigated a new method for the integrated detection and assembly of insertion variants from re-sequencing data. Contrary to other tools, it is designed to call insertions of any size, whether they are novel or duplicated, homozygous or heterozygous in the donor genome. We uses an efficient k-mer based method to detect insertion sites in a reference genome, and subsequently assemble them from the complete set of donor reads. The method is implemented in the tool MindTheGap and showed high recall and precision on simulated datasets of various genome complexities. When applied to real *C. elegans* and human datasets, MindTheGap detected and correctly assembled insertions longer than 1 kb, using at most 14 GB of memory. [19], [40]

Enhancement of de-Bruijn Graph data structure

The data structure holding the de-Bruijn Graph at the core of the GATB library has been improved through several new developments. First, its construction time has been greatly decreased thanks to the use of minimizers for kmer-counting, and efficient parallelization of various construction steps. Secondly, exploration of the graph has also been made faster through the possibility of parallel enumeration of nodes of interest, and through the use of a cache-coherent (blocked) bloom filter. Lastly, the structure itself has been extended to optionally allow for more information to be held, at a reasonable memory cost. A minimal perfect hash function allows to store additional data for each node, for example the coverage of each kmer. [11], [35], [36]

Chloroplast assembly

When sequencing plants, reads that correspond to the chloroplast genome are often over-represented. Filtering these reads based on k-mer counts allows specific assembly of the chloroplast to be directly performed. The small number of contigs can then be processed using advanced optimization tools to generate scaffolds. The approach has been partially tested on seqencing data from *Lactococcus lactis* to assemble plasmids of this bacteria. [12]

6.3. NGS applications

Participants: Susete Alves Carvalho, Rumen Andonov, Anaïs Gouin, Fabrice Legeai, Dominique Lavenier, Claire Lemaitre, Pierre Peterlongo, Ivaylo Petrov, Guillaume Rizk.

Identification of genomic regions of biological interest

The extraction and selection of 400 microsatellites among the large and fragmented *Acyrthosiphon pisum* genome led to the identification of a single 9cM region controlling the loss of sex in the pea aphid. The genotyping of these markers on geographically distant populations under divergent selection for reproductive strategies revealed a strong signature of selection in this genomic region, suggesting gene flow between populations with distinct reproductive modes. [15]

Transcriptome assembly

For this study, we incorporated various sources of RNA sequences from 454, Illumina and Sanger sequencing technologies obtained from more than 35 *S. frugiperda* developmental time-points and tissue samples and developed a custom pipeline to achieve their assembly. As a result, we provided a first valid transcriptome for *Spodoptera frugiperda*, a major agricultural pest. [16]

Catalogue of long non coding RNAs

We established a new bioinformatics pipeline for the detection of lncRNAs from RNA-Seq data, produced the first catalogue of aphids lncRNAs, and asserted for each lncRNA a classification of putative cis-interactions based on its genomic distance to neighboring mRNAs. These results allow the constitution of a broad gene regulation network of the aphid phenotypic plasticity at the embryo level. This workflow is available in Galaxy on the BioInformatics Platform for Arthopods of Agroecosystems (www.inra.fr/bipaa) and can be applied to any organism for which an annotated genome sequence and RNA-Seq data are provided.[23]

Identification and correction of genome mis-assemblies due to heterozygosity

Assembly tools are more and more efficient to reconstruct a genome from next-generation sequencing data but some problems remain. One of them corresponds to mis-assemblies due to heterozygosity (2 alleles instead of a consensus). Thus, we propose a strategy to detect and correct false duplications in assemblies based on several metrics: sequence similarity, matche length and average read coverage. Our method allows to decrease redundancy in the genome assembly, to improve the scaffolding and then to increase the N50 statistic by removal of one of the two alleles or joining of scaffolds by their extremities. This method was applied on the *Spodoptera frugiperda* genome.[39]

Questioning the classical re-sequencing analyses approach

Classical re-sequencing analyses are based on a first step of read mapping, then only mapped reads are taken into account in following analyses such as variant calling. We investigated the sources of unmapped reads in aphid re-sequencing data of 33 individuals, and we demonstrated that these reads contain valuable information that should not be discarded as usually done in such analyses. For instance, the analysis of the contigs obtained from assembling the unmapped reads led to recover some divergent genomic regions previously excluded from analysis and to discover putative novel sequences of *A. pisum* and its symbionts. We proposed strategies, based on assembly and re-mapping, to aid the capture and interpretation of this information.[14]

Application of discoSnp on pea data
The pea is a non-model organism with a large (4.5 GB) and complex genome which has not been sequenced yet. We compared, on the same set of low depth pea sequences, the SNPs generated by discoSnp with those obtained with a previous SNP discovery pipeline, and those generated using classical mapping approach combining Bowtie2 and GATK tools. [31]

6.4. HPC and parallelism

Participants: Rumen Andonov, Charles Deltel, Dominique Lavenier, François Moreews, Ivaylo Petrov.

Workflows
New tools are needed to enable the quick design and the intensive parallel execution of bioinformatics processes. Therefore, we propose a new Dataflow oriented workflow management system dedicated to intensive bioinformatics tasks. We worked on the interoperability of bioinformatics workflows using a model driven approach. Our results enable new import / export capabilities between multiple workflow management environments and incite to create a unique shared workflow model.[28]

Graph processing : the All-Pairs Shortest Paths problem
This research work anticipates the need of processing huge graphs that are results of intensive genomic sequence comparison (bank to bank processing). We proposed a new algorithm for solving the all-pairs shortest-path problem for planar graphs and graphs with small separators that exploits the massive on-chip parallelism available in today's Graphics Processing Units (GPUs). Our algorithm, based on the Floyd-War shall algorithm, has near optimal complexity in terms of the total number of operations, while its matrix-based structure is regular enough to allow for efficient parallel implementation on the GPUs. By applying a divide-and-conquer approach, we are able to make use of multi-node GPU clusters, resulting in more than an order of magnitude speedup over the fastest known Dijkstra-based GPU implementation and a two-fold speedup over a parallel Dijkstra-based CPU implementation.[27]

Benchmark of Alignment Search Tools
Comparing sequences is a daily task in bioinformatics and many software try to fulfill this need by proposing fast execution times and accurate results. Introducing a new software in this field requires to compare it to recognized tools with the help of well defined metrics. A set of quality metrics is proposed that enables a systematic approach for comparing alignment tools. These metrics have been implemented in a dedicated software, allowing to produce textual and graphical benchmark artifacts. [21]

6.5. Protein Structure

Participants: Rumen Andonov, Douglas Goncalves, Dominique Lavenier, Mathilde Le Boudic-Jamin, Antonio Mucherino.

The molecular distance geometry problem
The distance geometry is the problem of finding an embedding of a simple weighted undirected graph $G = (V, E, d)$ in a given dimension $K > 0$. Its most interesting application arises in biology, where the conformation of molecules such as proteins can be identified by embedding a graph (representing the molecular structure and some distance information) in dimension 3. Since some years, we are working on the discretization of the distance geometry. This year, the research developed in 4 main directions, that will be briefly detailed in the following paragraphs.

The majority of the work was performed on the so-called *discretization orders*, which are particular orders for the atoms of a molecule that allow for satisfying the discretization assumptions, i.e. they allow to discretize the search domain of the problem. Finding discretization orders is therefore an important pre-processing step for the solution of distance geometry problems. In fact, not only the identification of an atomic order allowing for the discretization is important, but also the identification of orders that are able to optimize some objectives that make the solution to the problem easier to perform. In this context, with both international and local partners, we worked on discretization orders that can be identified automatically in polynomial time [13], we worked on suitable orders for the protein side chains [10], and we studied some objectives to be optimized in discretization orders [38].

The algorithm that we mostly employ for the solution of distance geometry problems that can be discretized is the Branch & Prune (BP) algorithm. It recursively constructs the discretized search domain (a tree) and verifies the feasibility of the computed atomic positions. When all available distances are exact, all candidate positions for a given atom can be enumerated. This is however not possible in presence of interval distances, because a continuous subset of positions can actually be computed for the corresponding atoms. The focus of the work in [22] is on a new scheme for an adaptive generation of a discrete subset of candidate positions from this continuous subset. The generated candidate positions do not only satisfy the distances employed in the discretization process, but also additional distances that might be available (the so-called pruning distances).

Since the BP algorithm can loose in performance when dealing with large molecules containing several interval distances, we worked this year on a variation of the algorithm named BetaMDGP [29]. This is a work in collaboration with Korean researchers. The BetaMDGP algorithm is based on the concept of beta-complex, which is a geometric construct extracted from the quasi-triangulation derived from the Voronoi diagram of atoms.

From the theoretical side, we worked on two main directions. First, we proved that, in discretizable distance geometry problems where all available distances are exact, the total number of solutions is always a power of two. This is related to the fact that the discrete search space contains several symmetries [18]. Secondly, we tried to summarize in [37] the current issues for efficiently solving real-life instances of the distance geometry.

Finally, the work we performed during the last years, including another important results from other colleagues currently working on this topic, was summarized in an extensive survey on the discretization of the distance geometry [17].

Distance measure between Protein structure

We propose here a new distance measure for comparing two protein structures based on their contact map representations (CMO). This novel measure (max-CMO metric), satisfies all properties of a metric on the space of protein representations. Having a metric in that space allows to avoid pairwise comparisons on the entire database and thus to significantly accelerate exploring the protein space compared to non metric spaces. We show on a gold-standard classification benchmark sets that our exact k-nearest neighbor scheme classifies up to 95% and 99% of queries correctly. Our k-NN classification thus provides a promising approach for the automatic classification of protein structures based on contact map overlap. [26], [30]

Local similarity of protein structure

Finding similarities between protein structures is a main goal in molecular biology. Most of the existing tools preserve order and only find single alignments even when multiple similar regions exist. We propose a new seed-based approach that discovers multiple pairs of similar regions. Its computational complexity is polynomial and it comes with a quality guarantee that the returned alignments have both Root Mean Squared Deviations (coordinate-based as well as internal-distances based) lower than a given threshold, if such exists. We do not require the alignments to be order preserving, which makes our algorithm suitable for detecting similar domains when comparing multi-domain proteins. And because the search space for non-sequential alignments is much larger than for sequential ones, the computational burden is addressed by using both a coarse-grain level parallelism and a fine-grain level parallelism. [33]

GEOMETRICA Project-Team

6. New Results

6.1. Highlights of the Year

[10] was elected among the notable articles of 2013 by ACM and Computing Reviews (see http://computingreviews.com/recommend/bestof/notableitems_2013.cfm).

6.2. Mesh Generation and Geometry Processing

6.2.1. A Surface Reconstruction Method for In-Detail Underwater 3D Optical Mapping

Participant: Mariette Yvinec.

In collaboration with Pierre Alliez (EPI Titane), Ricard Campos (University of Girona), Raphael Garcia (University of Girona)

Underwater range scanning techniques are starting to gain interest in underwater exploration, providing new tools to represent the seafloor. These scans (often) acquired by underwater robots usually result in an unstructured point cloud, but given the common downward-looking or forward-looking configuration of these sensors with respect to the scene, the problem of recovering a piecewise linear approximation representing the scene is normally solved by approximating these 3D points using a heightmap (2.5D). Nevertheless, this representation is not able to correctly represent complex structures, especially those presenting arbitrary concavities normally exhibited in underwater objects. We present a method devoted to full 3D surface reconstruction that does not assume any specific sensor configuration. The method presented is robust to common defects in raw scanned data such as outliers and noise often present in extreme environments such as underwater, both for sonar and optical surveys. Moreover, the proposed method does not need a manual preprocessing step. It is also generic as it does not need any information other than the points themselves to work. This property leads to its wide application to any kind of range scanning technologies and we demonstrate its versatility by using it on synthetic data, controlled laser-scans, and multibeam sonar surveys. Finally, and given the unbeatable level of detail that optical methods can provide, we analyze the application of this method on optical datasets related to biology, geology and archeology. [23]

6.2.2. A Transfer Principle and Applications to Eigenvalue Estimates for Graphs

Participant: David Cohen-Steiner.

In collaboration with Omid Amini (ENS),

In this paper, we prove a variant of the Burger-Brooks transfer principle which, combined with recent eigenvalue bounds for surfaces, allows to obtain upper bounds on the eigenvalues of graphs as a function of their genus. More precisely, we show the existence of a universal constants C such that the k-th eigenvalue λ_k of the normalized Laplacian of a graph G of (geometric) genus g on n vertices satisfies $\lambda_k \leq C d_{max}(g+k)/n$ where $d_m ax$ denotes the maximum valence of vertices of the graph. This result is tight up to a change in the value of the constant C. We also use our transfer theorem to relate eigenvalues of the Laplacian on a metric graph to the eigenvalues of its simple graph models, and discuss an application to the mesh partitioning problem. [44]

6.3. Topological and Geometric Inference

6.3.1. Only distances are required to reconstruct submanifolds

Participants: Jean-Daniel Boissonnat, Steve Oudot.

In collaboration with Ramsay Dyer (Johann Bernouilli Institute, University of Groningen, Pays Bas) and Arijit Ghosh (Max-Planck-Institut für Informatik, Saarbrücken, Germany).

In [45], we give the first algorithm that outputs a faithful reconstruction of a submanifold of Euclidean space without maintaining or even constructing complicated data structures such as Voronoi diagrams or Delaunay complexes. Our algorithm uses the witness complex and relies on the stability of *power protection*, a notion introduced in this paper. The complexity of the algorithm depends exponentially on the intrinsic dimension of the manifold, rather than the dimension of ambient space, and linearly on the dimension of the ambient space. Another interesting feature of this work is that no explicit coordinates of the points in the point sample is needed. The algorithm only needs the *distance matrix* as input, i.e., only distance between points in the point sample as input.

6.3.2. Computing Persistent Homology with Various Coefficient Fields in a Single Pass

Participants: Jean-Daniel Boissonnat, Clément Maria.

In [32], we introduce an algorithm to compute the persistent homology of a filtered complex with various coefficient fields in a single matrix reduction. The algorithm is output-sensitive in the total number of distinct persistent homological features in the diagrams for the different coefficient fields. This computation allows us to infer the prime divisors of the torsion coefficients of the integral homology groups of the topological space at any scale, hence furnishing a more informative description of topology than persistence in a single coefficient field. We provide theoretical complexity analysis as well as detailed experimental results.

6.3.3. Recognizing shrinkable complexes is NP-complete

Participants: Olivier Devillers, Marc Glisse.

In collaboration with Dominique Attali (Gipsa-lab, Grenoble), Sylvain Lazard (Inria Nancy - Grand Est)

We say that a simplicial complex is shrinkable if there exists a sequence of admissible edge contractions that reduces the complex to a single vertex. We prove [31] that it is NP-complete to decide whether a (three-dimensional) simplicial complex is shrinkable. Along the way, we describe examples of contractible complexes which are not shrinkable.

6.3.4. Zigzag Zoology: Rips Zigzags for Homology Inference

Participant: Steve Oudot.

In collaboration with Donald Sheehy (University of Connecticut)

For points sampled near a compact set X, the persistence barcode of the Rips filtration built from the sample contains information about the homology of X as long as X satisfies some geometric assumptions. The Rips filtration is prohibitively large, however zigzag persistence can be used to keep the size linear. We present [28] several species of Rips-like zigzags and compare them with respect to the signal-to-noise ratio, a measure of how well the underlying homology is represented in the persistence barcode relative to the noise in the barcode at the relevant scales. Some of these Rips-like zigzags have been available as part of the Dionysus library for several years while others are new. Interestingly, we show that some species of Rips zigzags will exhibit less noise than the (non-zigzag) Rips filtration itself. Thus, the Rips zigzag can offer improvements in both size complexity and signal-to-noise ratio. Along the way, we develop new techniques for manipulating and comparing persistence barcodes from zigzag modules. We give methods for reversing arrows and removing spaces from a zigzag. We also discuss factoring zigzags and a kind of interleaving of two zigzags that allows their barcodes to be compared. These techniques were developed to provide our theoretical analysis of the signal-to-noise ratio of Rips-like zigzags, but they are of independent interest as they apply to zigzag modules generally.

6.3.5. Zigzag Persistence via Reflections and Transpositions

Participants: Clément Maria, Steve Oudot.

We introduce [40] a simple algorithm for computing zigzag persistence, designed in the same spirit as the standard persistence algorithm. Our algorithm reduces a single matrix, maintains an explicit set of chains encoding the persistent homology of the current zigzag, and updates it under simplex insertions and removals. The total worst-case running time matches the usual cubic bound. A noticeable difference with the standard persistence algorithm is that we do not insert or remove new simplices "at the end" of the zigzag, but rather "in the middle". To do so, we use arrow reflections and transpositions, in the same spirit as reflection functors in quiver theory. Our analysis introduces a new kind of reflection called the "weak-diamond", for which we are able to predict the changes in the interval decomposition and associated compatible bases. Arrow transpositions have been studied previously in the context of standard persistent homology, and we extend the study to the context of zigzag persistence. For both types of transformations, we provide simple procedures to update the interval decomposition and associated compatible homology basis.

6.3.6. Topological analysis of scalar fields with outliers

Participants: Mickaël Buchet, Frédéric Chazal, Steve Oudot.

In collaboration with Tamal K. Dey (University of Ohio) Fengtao Fan (University of Ohio) Yusu Wang (University of Ohio)

We extend [57] the notion of the distance to a measure from Euclidean space to probability measures on general metric spaces as a way to do topological data analysis in a way that is robust to noise and outliers. We then give an efficient way to approximate the sub-level sets of this function by a union of metric balls and extend previous results on sparse Rips filtrations to this setting. This robust and efficient approach to topological data analysis is illustrated with several examples from an implementation.

6.3.7. Efficient and Robust Persistent Homology for Measures.

Participants: Mickaël Buchet, Frédéric Chazal, Steve Oudot.

In collaboration with Donald Sheehy (University of Connecticut)

In [34], we extend the notion of the distance to a measure from Euclidean space to probability measures on general metric spaces as a way to do topological data analysis in a way that is robust to noise and outliers. We then give an efficient way to approximate the sub-level sets of this function by a union of metric balls and extend previous results on sparse Rips filtrations to this setting. This robust and efficient approach to topological data analysis is illustrated with several examples from an implementation.

6.3.8. Persistence-based Structural Recognition

Participants: Frédéric Chazal, Maksims Ovsjanikovs.

In collaboration with Chunyuan Li (former intern in Saclay in 2013)

In [39] we present a framework for object recognition using topological persistence. In particular, we show that the so-called persistence diagrams built from functions defined on the objects can serve as compact and informative descriptors for images and shapes. Complementary to the bag-of-features representation, which captures the distribution of values of a given function, persistence diagrams can be used to characterize its structural properties, reflecting spatial information in an invariant way. In practice, the choice of function is simple: each dimension of the feature vector can be viewed as a function. The proposed method is general: it can work on various multimedia data, including 2D shapes, textures and triangle meshes. Extensive experiments on 3D shape retrieval, hand gesture recognition and texture classification demonstrate the performance of the proposed method in comparison with state-of-the-art methods. Additionally, our approach yields higher recognition accuracy when used in conjunction with the bag-offeatures.

6.3.9. Convergence rates for persistence diagram estimation in Topological Data Analysis

Participants: Frédéric Chazal, Marc Glisse, Bertrand Michel.

In collaboration with Catherine Labruère (University of Burgundy)

Computational topology has recently known an important development toward data analysis, giving birth to the field of topological data analysis. Topological persistence, or persistent homology, appears as a fundamental tool in this field. In [36], we study topological persistence in general metric spaces, with a statistical approach. We show that the use of persistent homology can be naturally considered in general statistical frameworks and persistence diagrams can be used as statistics with interesting convergence properties. Some numerical experiments are performed in various contexts to illustrate our results.

6.3.10. Stochastic Convergence of Persistence Landscapes and Silhouettes

Participant: Frédéric Chazal.

In collaboration with Brittany Fasy (Tulane University) Fabrizio Lecci (Carnegie Mellon University) Alessandro Rinaldo (Carnegie Mellon University) Larry Wasserman (Carnegie Mellon University)

Persistent homology is a widely used tool in Topological Data Analysis that encodes multiscale topological information as a multi-set of points in the plane called a persistence diagram. It is difficult to apply statistical theory directly to a random sample of diagrams. Instead, we can summarize the persistent homology with the persistence landscape, introduced by Bubenik, which converts a diagram into a well-behaved real-valued function. In [35], we investigate the statistical properties of landscapes, such as weak convergence of the average landscapes and convergence of the bootstrap. In addition, we introduce an alternate functional summary of persistent homology, which we call the silhouette, and derive an analogous statistical theory.

6.3.11. Subsampling Methods for Persistent Homology

Participants: Frédéric Chazal, Bertrand Michel.

In collaboration with Brittany Fasy (Tulane University) Fabrizio Lecci (Carnegie Mellon University) Alessandro Rinaldo (Carnegie Mellon University) Larry Wasserman (Carnegie Mellon University)

Persistent homology is a multiscale method for analyzing the shape of sets and functions from point cloud data arising from an unknown distribution supported on those sets. When the size of the sample is large, direct computation of the persistent homology is prohibitive due to the combinatorial nature of the existing algorithms. We propose to compute the persistent homology of several subsamples of the data and then combine the resulting estimates. We study the risk of two estimators and we prove that the subsampling approach carries stable topological information while achieving a great reduction in computational complexity.

6.3.12. The observable structure of persistence modules

Participant: Frédéric Chazal.

In collaboration with Vin de Silva (Pomona College) William Crawley-Boevey (University of Leeds)

In persistent topology, q-tame modules appear as a natural and large class of persistence modules indexed over the real line for which a persistence diagram is de- finable. However, unlike persistence modules indexed over a totally ordered finite set or the natural numbers, such diagrams do not provide a complete invariant of q-tame modules. The purpose of [59] is to show that the category of persistence modules can be adjusted to overcome this issue. We introduce the observable category of persis- tence modules: a localization of the usual category, in which the classical properties of q-tame modules still hold but where the persistence diagram is a complete isomorphism invariant and all q-tame modules admit an interval decomposition.

6.4. Data Structures and Robust Geometric Computation

6.4.1. Efficiently Navigating a Random Delaunay Triangulation

Participants: Olivier Devillers, Ross Hemsley.

In collaboration with Nicolas Broutin (EPI RAP)

Planar graph navigation is an important problem with significant implications to both point location in geometric data structures and routing in networks. Whilst many algorithms have been proposed, very little theoretical analysis is available for the properties of the paths generated or the computational resources required to generate them. In this work, we propose and analyse a new planar navigation algorithm for the Delaunay triangulation. We then demonstrate a number of strong theoretical guarantees for the algorithm when it is applied to a random set of points in a convex region [33]. In a side result, we give a new polylogarithmic bound on the maximum degree of a random Delaunay triangulation in a smooth convex, that holds with probability one as the number of points goes to infinity. In particular, our new bound holds even for points arbitrarily close to the boundary of the domain. [56]

6.4.2. A chaotic random convex hull

Participants: Olivier Devillers, Marc Glisse, Rémy Thomasse.

The asymptotic behavior of the expected size of the convex hull of uniformly random points in a convex body in $\mathcal{R}^d$ is polynomial for a smooth body and polylogarithmic for a polytope. We construct a body whose expected size of the convex hull oscillates between these two behaviors when the number of points increases [62]

6.4.3. A generator of random convex polygons in a disc

Participants: Olivier Devillers, Rémy Thomasse.

In collaboration with Philippe Duchon (LABRI)

Let $\mathcal{D}$ a disc in $\mathbb{R}^2$ with radius 1 centered at $\mathfrak{o}$, and $(x_1, \cdots, x_n)$ a sample of n points uniformly and independently distributed in $\mathcal{D}$. Let's define the polygon P_n as the convex hull of $(x_1, \cdots, x_n)$, and $f_0(P_n)$ its number of vertices. This kind of polygon has been well studied, and it is known, see [65], that

$$\mathbb{E}f_0(P_n) = c\ n^{\frac{1}{3}} + o(n^{\frac{1}{3}})$$

where $c > 0$ is constant. To generate such a polygon, one can explicitly generate n points uniformly in $\mathcal{D}$ and compute the convex hull. For a very large quantity of points, it could be interesting to generate less points to get the same polygon, for example to have some estimations on asymptotic properties, such as the distribution of the size of the edges.We propose an algorithm that generate far less points at random in order to get P_n, so that the time and the memory needed is reduced for n large. Namely [61], we generate a number of points of the same order of magnitude than the final hull, up to a polylogarithmic factor

6.4.4. On the complexity of the representation of simplicial complexes by trees

Participants: Jean-Daniel Boissonnat, Dorian Mazauric.

In [46], we investigate the problem of the representation of simplicial complexes by trees. We introduce and analyze local and global tree representations. We prove that the global tree representation is more efficient in terms of time complexity for searching a given simplex and we show that the local tree representation is more ecient in terms of size of the structure. The simplicial complexes are modeled by hypergraphs. We then prove that the associated combinatorial optimization problems are very dicult to solve and to approximate even if the set of maximal simplices induces a cubic graph, a planar graph, or a bounded degree hypergraph. However, we prove polynomial time algorithms that compute constant factor approximations and optimal solutions for some classes of instances.

6.4.5. Building Efficient and Compact Data Structures for Simplicial Complexes

Participant: Jean-Daniel Boissonnat.

In collaboration with Karthik C.S (Weizmann Institute of Science, Israël) and Sébastien Tavenas (Max-Planck-Institut für Informatik, Saarbrücken, Germany).

The Simplex Tree is a recently introduced data structure that can represent abstract simplicial complexes of any dimension and allows to efficiently implement a large range of basic operations on simplicial complexes. In this paper, we show how to optimally compress the simplex tree while retaining its functionalities. In addition, we propose two new data structures called Maximal Simplex Tree and Compact Simplex Tree. We analyze the Compressed Simplex Tree, the Maximal Simplex Tree and the Compact Simplex Tree under various settings.

6.4.6. Delaunay triangulations over finite universes

Participant: Jean-Daniel Boissonnat.

In collaboration with Ramsay Dyer (Johann Bernouilli Institute, University of Groningen, Pays Bas) and Arijit Ghosh (Max-Planck-Institut für Informatik, Saarbrücken, Germany).

The witness complex was introduced by Carlsson and de Silva as a weak form of the Delaunay complex that is suitable for finite metric spaces and is computed using only distance comparisons. The witness complex $\mathrm{Wit}(L, W)$ is defined from two sets L and W in some metric space X: a finite set of points L on which the complex is built, and a set W of witnesses that serves as an approximation of X. A fundamental result of de Silva states that $\mathrm{Wit}(L, W) = \mathrm{Del}(L)$ if $W = X = \mathbb{R}^d$. In this paper we give conditions on L that ensure that the witness complex and the Delaunay triangulation coincide when $W \subset \mathbb{R}^d$ is a finite set, and we introduce a new perturbation scheme to compute a perturbed set L' close to L such that $\mathrm{Del}(L') = \mathrm{Wit}(L', W)$. The algorithm constructs $\mathrm{Wit}(L', W)$ in time sublinear in $|W|$.

The only numerical operations used by our algorithms are (squared) distance comparisons (i.e., predicates of degree 2). In particular, we do not use orientation or in-sphere predicates, whose degree depends on the dimension d, and are difficult to implement robustly in higher dimensions. Although the algorithm does not compute any measure of simplex quality, a lower bound on the thickness of the output simplices can be guaranteed. Another novelty in the analysis is the use of the Moser-Tardos constructive proof of the general Lovász local lemma.

GEOSTAT Project-Team

6. New Results

6.1. Highlights of the Year

Paper **Spanning the Scales of Granular Materials through Microscopic Force Imaging** by N. Brodu *et al.* accepted in **Nature Communications** (will appear in 2015).
BEST PAPER AWARD :
[36] **IEEE TENSYMP 2014**. B. XU, S. BINCZAK, S. JACQUIR, O. PONT, H. YAHIA.

6.2. Super-resolution for Earth Observation data

Participants: Hussein Yahia, Joël Sudre, Oriol Pont, Véronique Garçon, Dharmendra Singh.

References: [17], [30], [28], [38], [29].
With partners at LEGOS and in the framework of the OPTIC associated team (7.4.1), we are developping the novel super-resolution approaches for Universe Sciences data. New results are obtained for ocean dynamics, partial pressures pCO_2 between the ocean and the atmosphere, and data fusion.

6.3. Fast and Accurate Texture Recognition with Multilayer Convolution and Multifractal Analysis

Participants: Hicham Badri, Hussein Yahia, Khalid Daoudi.

Reference: [25].
A fast and accurate texture recognition system is presented. It consists in extracting locally and globally invariant representations of a given texture image. The mapping from the locally to the globally invariant representation is based on a scale-invariant method via the calculation of singularity exponents. The final descriptor is extracted from the distribution of these exponents and leads to a more accurate descriptor compared to the popular box-counting method. We also propose to use a combination of the generative PCA classifier together with multi-class SVM as well as a synthetic training strategy. Experiments show that the proposed solution outperforms existing methods on three challenging public benchmark datasets, while being computationally efficient.

6.4. Fast Image Edge-Aware Processing

Participants: Hicham Badri, Hussein Yahia, Driss Aboutajdine.

Reference: Article *Fast Edge-Aware Processing via First Order Proximal Approximation* by H. Badri, H. Yahia, D. Aboutajdine, accepted with minor revision in **IEEE Transactions on Visualization & Computer Graphics**, will be in HAL in 2015.
We present a framework for fast edge-aware processing of images and videos. This is an extension of our previous SIGGRAPH Asia 2013 paper. The proposed approach uses non-convex sparsity on the gradients of the latent smooth image to better preserve sharp edges. We develop tools based on first order proximal estimation for fast processing. We also propose fast and efficient numerical solutions based on separable filters estimation, which enables our method to perform fast high-quality smoothing on large-scale images. Extensive experiments show that the proposed method produces high-quality smoothing compared to state-of-the-art methods, while being fast and simple to implement.

6.5. Cardiac arrhythmia induced by mild hypothermia in vitro – a pitchfork bifurcation type process

Participants: Binbin Xu, Oriol Pont.

Reference: [20].
The neurological damage after cardiac arrest constitutes a big challenge of hospital discharge. The mild therapeutic hypothermia (MTH) (34°C - 32°C) has shown its benefit to reduce this type of damage. However, it can have many adverse effects, among which the cardiac-arrhythmia-generation-a-posteriori (CAGP) can represent up to 34%. Our study with a cardiac culture in vitro showed that at 35°C the CAGP can be induced. The process of MTH can be represented by a Pitchfork bifurcation, which could explain the different ratio of arrhythmia among the adverse effects after this therapy. This nonlinear dynamics suggests that a variable speed of cooling / rewarming, especially when passing 35°C, would help to decrease the ratio of post-hypothermia arrhythmia and then improve the hospital output.

6.6. Characterizing the dynamics of cardiac arrhythmia

The dynamics of cardiac arrhythmia is quite complex. Better understanding its mechanism can help to improve the treatment. In vitro cultures of cardiac cells which has similar parameters as cell of human's heart represent valuable tool and model to study this issue.

6.6.1. by Complexity Analysis

Participants: Binbin Xu, Oriol Pont.

References: [36], [39].
Stochastic approaches provide a type of methods to characterize cardiac arrhythmia, aimed at quantifying the statistical properties of the time series. Complexity analysis such as Approximate Entropy (ApEn) and Sample Entropy (SampEn), are particularly useful to analyze time series in electro-cardiology in which the signals are characterized by their high regularity in normal condition in contrast to irregularity in pathological cases. It is shown that ApEn and SampEn can not only serve as a discrimination index, but also provide another parameter which showed doubling phenomenon. It proves in other terms that bifurcation happens in case arrhythmia. See figure 1 .

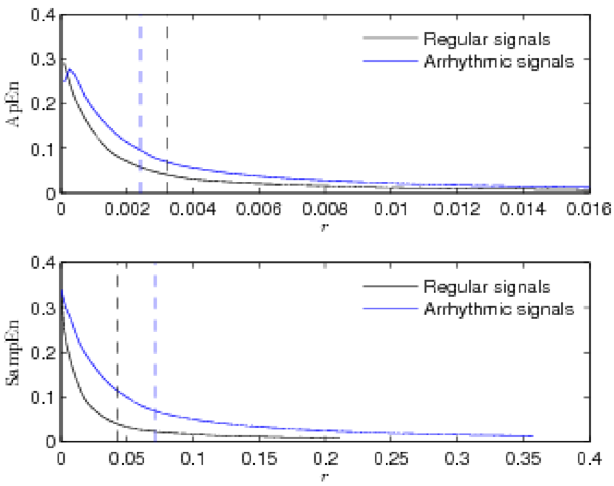


Figure 1. Illustration of ApEn / SampEn analysis for normal and arrhythmic electrical field potential.

6.6.2. by Phase Space Reconstruction

Participants: Oriol Pont, Binbin Xu.

References: [19], [40], [22].
Phase space reconstructions of electrical field potential signals in normal and arrhythmic cases are performed by characterizing the nonlinearity of these signals. The phase space reconstructions highlight attractors, whose dimension reveals that they are strange, depicting a deterministic dynamics of chaotic nature in the in vitro model. The electrical activity of the heart consists of nonlinear interactions emerging as a complex system. Electrocardiographic imaging provides a full spatiotemporal picture of the electric potential on the human epicardium. Rhythm reflects the connection topology of the pacemaker cells driving it. Hence, characterizing the attractors as nonlinear, effective dynamics can capture the key parameters without imposing any particular microscopic model on the empirical signals. A dynamic phase-space reconstruction from an appropriate embedding can be made robust and numerically stable with the methods developed in the team. With these, we have been able to show how both the phase-space descriptors and those of the a priori unrelated singularity analysis are able to highlight the arrhythmogenic areas on cases of atrial fibrillation. See figure 2 .

Figure 2. Illustration of phase space analysis of normal and arrhythmic electrical field potential.

6.7. The origin of the myth FitzHugh-Nagumo model

Participants: Binbin Xu, Oriol Pont.

Reference: [37].
History became legend. Legend became myth. Derived from the pioneer ionic Hodgkin-Huxley model and due to its simplicity and richness from a point view of nonlinear dynamics, the FitzHugh-Nagumo model (FHN) is one of the most successful simplified neuron / cardiac cell model. 60 years later, there exist many variations of this model whose parameters (ε, γ and α) are often used in biased conditions. The related results would be questionable. This study showed that α controls the global dynamics of FHN. $\alpha > 0$, the cell is in refractory mode and does not respond to external stimulation; if $\alpha < 0$, the cell is excitable. ε controls the main morphology of the action potential generated. γ influences barely AP, it showed linear relationship with the period and duration of AP. Though it can be freely chosen for excitable cell, but smaller values are recommended.

6.8. Pathological Speech Analysis

Participants: Khalid Daoudi, Vahid Khanagha, Blaise Bertrac, Safa Mrad, Ashwini Jaya Kumar.

References: [14], [13], [26], [27].

We applied our recent results in nonlinear speech analysis to the filed of pathological speech detection and classification. We presented new insights in the task of normal-vs-pathological voice classification using the widely used Kayelemetrics database. In particular, we showed that hat one single parameter, derived from matching pursuit decomposition of speech, allow perfect discrimination between normal and dysphonic voices of these database. This result raises some important questions on the way this task is generally addressed. Using our GCI detection algorithm, we also proposed new definitions of standard voice perturbation measures (jitter, shimmer...) which lead to significantly higher classification accuracy. Our new measures have the strong advantage to avoid the usual periodicity and linearity assumptions. On the other hand, we started investigating the task of discrimination between Parkinson's and healthy voices. Our phonetic segmentation algorithm has potentially the ability to detect vowel onset and offset regions which have different structures in Parkinson's voices that in healthy ones. This preliminary result is promising and we are continuing research in this direction.

6.9. Statistics and detection of most unpredictable points in data sets

Participants: Nicolas Brodu, Hussein Yahia, Suman-Kumar Maji.

References: [21], [16].

The assumption that local regularity amounts to predictability can be challenged, depending on the model that one may use to make predictions. A statistical framework, "computational mechanics", has been explicitly designed over the past 30 years, that precisely formalizes notions of causality and predictability within discrete data sets. Patterns with similar causal influence on the data are clustered in equivalence classes. Taken together, these classes form a Markovian automaton by definition, since no extra information is needed from other classes to (statistically) predict the influence of a group of patterns on the rest of the data set. These automata are defined at the lowest data description scale, but it has been suggested that sub-automata (thus clusters at larger scales) form an ideal coarse-graining of the system in terms of predictability (thus also descriptive power). The theory is also deeply rooted in statistical physics, offering a unique perspective on how macroscopic variables could be derived from a microscopic description of a studied system. Preliminary results are promising and show that, for example, edges may be detected in images with a precursor continuous implementation of the theory extension under construction. In order to make more progress, advanced statistical and computational developments are necessary to carry this work. In order to facilitate this development, N. Brodu has submitted a Marie-Curie outgoing fellowship that, if accepted, would allow to partner with Australian leaders on statistics and data processing (University of Melbourne, department of Mathematics).

6.10. Image Reconstruction from Highly Corrupted Gradients

Participants: Hicham Badri, Hussein Yahia, Driss Aboutajdine.

Reference: [23].

Surface-from-Gradients (SfG) is an important step in many imaging applications. It consists in reconstructing an image/surface from corrupted gradient fields, which results in an ill-posed problem. We propose to use sparsity to regularize the problem. We use sparsity in the gradient field together with a robust norm on the data-fitting term (CVPR 2014). In a work in porgress, we make use of a non-local regularization that manipulates non-local similar patches of the corrupted gradient and forcing them to be low-rank. The two approaches significantly outperform previous optimization-based SfG methods on both synthetic and real data.

6.11. Local/Non-Local Noisy Image Deconvolution

Participants: Hicham Badri, Hussein Yahia.

Reference: [24].

Image deconvolution is a standard step in many imaging applications. Sparse local regularization has shown to be fast but tends to over-smoothing images. On the other hand, non-local priors that manipulate similar patches produce better results but tend to be much slower. In this paper, we combine both local and non-local methods in one framework to offer both good quality image reconstruction and computational efficiency in the presence of noise. By studying the non-local singular values of the image patches, we show that the non-local patches tend to be much similar in the blurred version of the image. We thus use low-rank estimation to first estimate a blurred but noise-free image. Secondly, we show that this denoising step introduces outliers in the deconvotion model and propose anefficient optimization method to tackle this problem. Experiments show that the proposed method poduces comparable results to non-local methods while being more computationally efficient.

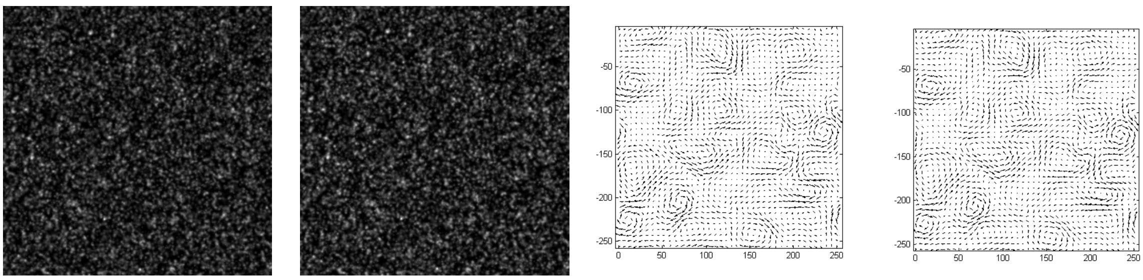


Figure 3. Motion estimation using the proposed method. From left to right: image sequences (2 images, at t and $t+1$ respectively) the ground-truth and the estimated flow (errors, from left to right : MSE=0.063, AAE=3.562, EPE=0.100).

6.12. Detection and dynamics of coastal upwelling

Participants: Ayoub Tamim, Khalid Daoudi, Hussein Yahia, Joël Sudre, Driss Aboutajdine.

References: [18], [34], [35], [33].
An unsupervised classification method is developed for the coarse segmentation of Moroccan coastal upwelling using the Sea Surface Temperature (SST) satellite images. The algorithm is used to provide a seasonal variability of upwelling activity in the southern Moroccan Atlantic coast using 70 Sea Surface Temperature (SST) images of the years 2007 and 2008. The performance of the proposed methodology has been validated by an oceanographer, showing its effectiveness for automatic delimitation of Moroccan upwelling region. We have also explored the applicability of the Fuzzy c-means (FCM) clustering, using an adaptive cluster merging, for the problem of detecting the Moroccan coastal upwelling areas in SST satellite images.

6.13. Nonlinear signal processing for adaptive optics

Participants: Suman-Kumar Maji, Hussein Yahia, Thierry Fusco.

Reference: [31].
The work developped by PhD student Suman Kumar Maji on nonlinear approaches to phase reconstruction in adaptive optics has been presented at the SPIE Astronomical Telescopes + Instrumentation, one of the great events in the field.

6.14. Turbulent Flow Estimation

Participants: Hicham Badri, Hussein Yahia.

We use singularity exponents (SE) to regularize the problem of turbulent flow estimation under the assumption that the brightness constancy constraint holds also for (SE). We also use weighted filtering (Lucas–Kanade's solution) and sparsity on the data-fitting term to improve robusteness to outliers. The proposed motion estimation is built on a Gaussian pyramid and uses the theory of warping for a better estimation of large displacements. Experiements on synthetic data show that the proposed method outperforms sophisticated methods while being simple. See figure 3 .

6.15. Adaptive Transfer Real Image Restoration

Participants: Hicham Badri, Hussein Yahia.

Image restoration is a very challenging task in low-level vision and is extensively used in many imaging applications. Sparsity in various forms (dictionary learning, low-rank estimation,...) has shown to be the key for succesful image denoising. However, the standard noise model used to validate the results is mainly Gaussian and uniform, with known standard deviation. Unfortunately, these assumptions do not hold for real camera noise. Instead of using sparsity to model the singular values of non-local clean similar patches, we use a learning model that trains a mapping between the noisy and ground-truth clean singular values. The training is performed on real camera noise, contrary to previous methods. Experiments show that the proposed method signficantly outperforms previous denoising works on real non-uniform noise and does not require estimating the standard deviation of the corruption. See figure 4 .

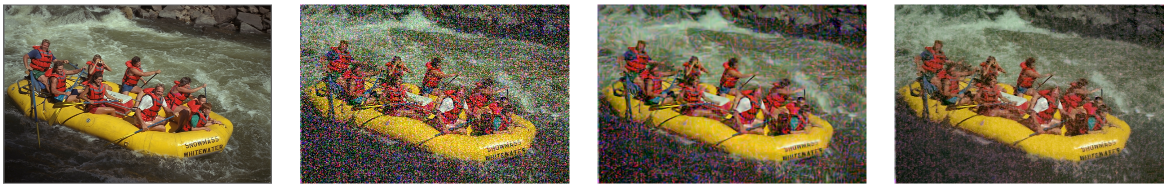


Figure 4. Image restoration demonstration on a severely corrupted image. The proposed method leads to a much better resotration quality compared to the standard BM3D method. From left to right: Ground-Truth, Noisy image, BM3D (20.46 dB), Proposed (22.25 dB).

6.16. Augmented Lagrangian for Fast Multi-Sparse Optimization

Participants: Hicham Badri, Hussein Yahia, Khalid Daoudi.

Sparsity has become one of the most important notions in many imaging applications. We address in this paper the problem of multi-sparse optimization, when the energy to minimize contains multiple sparse terms instead of a single one. We show that applying off-the-shelf proximal-based solvers such as ADMM results in a high computational cost due to the complexity of the resulting sub-problems in the case of multi-sparsity. We propose an efficient extension of ADMM for multi-sparse optimization, we study its convergence and complexity and show how it can be applied to computer vision problems. Experiments show that the proposed solver is not only computationally efficient, but also leads quickly to higher-quality results compared to the popular half-quadratic solver.

6.17. On the Fly Hybrid Video Denoising

Participants: Hicham Badri, Hussein Yahia.

Video denoising is a standard pre-processing step in many imaging applications. Non-local methods such as the BM3D method adapted to videos have shown to produce good quality results, but these methods require multiple frames to produce a temporally coherent result, especially when the amount of noise is high. On the other hand, using a hybrid camera, we can get clean images of the scene. However, these images suffer from low-temporal coherence. We present a new approach to video denoising which consists in learning a mapping between the clean images and their corresponding noisy frames and propagate denoising to intermediate frames. To improve temporal coherency, we use a fast method method to sparsify the temporal gradient. Experiments on high-resolution videos show that the proposed method produces good quality on the fly video denoising while being computationally efficient.

GRACE Project-Team

6. New Results

6.1. Highlights of the Year

- F. Morain and A. Guillevic (with their co-authors R. Barbulescu and P. Gaudry) broke the discrete logarithm world record for finite fields of the form $GF(p^2)$ with a prime p of 80 decimal digits. The new techniques form the preprint [31].
- D. Augot and M. Finiasz received the best paper award at FSE 2014 [17]. FSE is the most important conference devoted to symmetric cryptography. Grace contribution is to propose a mathematical construction which enables direct construction of so-called diffusion layers in block ciphers.
- A. Zeh, former Grace PhD student, received the special Prize of the Université Franco-Allemande (UFA) Jury 2014 at the French Ambassy in Berlin, on November 21st.

BEST PAPER AWARD :

[17] **21st International Workshop on Fast Software Encryption, FSE 2014**. D. AUGOT, M. FINIASZ.

6.2. Diffusion layers for block ciphers

MDS matrices allow the construction of optimal linear diffusion layers in block ciphers. However, MDS matrices usually have a large description (for example, they can never be sparse), and this results in costly software/hardware implementations. We can solve this problem using *recursive MDS matrices*, which can be computed as a power of a simple companion matrix—and thus have a compact description suitable for constrained environments. Until now, finding recursive MDS matrices required an exhaustive search on families of companion matrices; this clearly limited the size of MDS matrices that one could look for. We have found a new direct construction, based on shortened BCH codes, which allows us to efficiently construct these matrices for arbitrary parameter sizes [17]. D. Augot and M. Finiasz received the best paper award at FSE 2014, and were invited to submit an extended journal version to *Journal of Cryptology*.

P. Karpman started to study sub-optimal diffusion layers, which can be built using algebraic geometry codes with a large automorphism group. Preliminary work has been done, leading to promising results [18]. To properly assert the cryptanalytic properties of these codes, V. Ducet is starting to implement a method for computing efficiently the weight distribution of AG codes.

6.3. Rank metric codes over infinite fields

Rank metric and Gabidulin codes over the rationals promise interesting applications to space-time coding. We have constructed optimal codes, similar to Gabidulin codes, in the case of infinite fields. We use algebraic extensions, and we have determined the condition on the considered extension to enable this construction. For example: we can design codes with complex coefficients, using number fields and Galois automorphisms. Then, in the rank metric setting, codewords can be seen as matrices. In this setting, a channel introduces errors (a matrix of small rank r added to the codeword) and erasures (s_r rows and s_c columns of the matrix are erased). We have developed an algorithm (adapted from the Welch–Berlekamp algorithm) to recover the right codeword in the presence of an error of rank weight up to $r + s_c + s_r \leqslant d - 1$, where d is the minimal distance of the code. As opposed to the finite field case, we are confronted by coefficient size growth. We solve this problem by computing modulo prime ideals. Using these codes we can completely bypass intermediate constructions using finite fields, which were the stumbling-block in classic constructions.

We also have used this framework to build rank-metric codes over the field of rational functions, using algebraic function fields with cyclic Galois group (Kummer and Artin extensions). These codes can be seen as a generator of infinitely many convolutional codes [25].

6.4. Tensor rank of multiplication over finite fields

Determining the tensor rank of multiplication over finite fields is a problem of great interest in algebraic complexity theory, but it also has practical importance: it allows us to obtain multiplication algorithms with a low bilinear complexity, which are of crucial significance in cryptography. In collaboration with S. Ballet and J. Chaumine [35], J. Pieltant obtained new asymptotic bounds for the symmetric tensor rank of multiplication in finite extensions of finite fields $\mathbb{F}_q$. In the more general (not-necessarily-symmetric) case, J. Pieltant and H. Randriam obtained new uniform upper bounds for multiplication in extensions of $\mathbb{F}_q$. They also gave purely asymptotic bounds substantially improving those coming from uniform bounds, by using a family of Shimura curves defined over $\mathbb{F}_q$. This work will appear in Mathematics of Computation [15].

6.5. Filtration Attacks against McEliece Cryptosystem

The McEliece encryption scheme based on binary Goppa codes was one of the first public-key encryption schemes [39]. Its security rests on the difficulty of decoding an arbitrary code. The original proposal uses classical Goppa codes, and while it still remains unbroken, it requires a huge size of key. On the other hand, many derivative systems based on other families of algebraic codes have been subject to key recovery attacks. Up to now, key recovery attacks were based either on a variant of Sidelnikov and Shestakov's attack [40], where the first step involves the computation of minimum-weight codewords, or on the resolution of a system of polynomial equations using Gröbner bases.

In [10], A. Couvreur, P. Gaborit, V. Gauthier, A. Otmani and J.-P. Tillich introduced a new paradigm of attack called *filtration attacks*. The general principle decomposes in two steps:

1. **Distinguishing** the public code from a random one using the square code operation.
2. **Computing a filtration** of the public code using the distinguisher, and deriving from this filtration an efficient decoding algorithm for the public code.

This new style of attack allowed A. Couvreur, A. Otmani and J.-P. Tillich to break (in polynomial time) McEliece based on wild Goppa codes over quadratic extensions [23]; and A. Couvreur, I. Márquez-Corbella, and R. Pellikaan to break McEliece based on algebraic geometry codes from curves of arbitrary genus [22], [26].

6.6. A new bound on the number of rational points of arbitrary projective varieties

In [38], the authors asked for a general upper bound on the number of rational points of a (possibly reducible) equidimensional variety $X \subseteq \mathbf{P}^n$ of dimension d and degree δ. They conjectured that

$$|X(\mathbf{F}_q)| \leq \delta(\pi_d - \pi_{2d-n}) + \pi_{2d_n}, \qquad (6)$$

where for all positive integer ℓ, π_ℓ is defined as the number of rational points of the projective space of dimension ℓ over $\mathbf{F}_q$. That is to say, $\pi_\ell = \frac{q^{\ell+1}-1}{q-1}$.

By combining algebraic geometric methods with a combinatorial method of double counting, A. Couvreur proved this conjecture [32] and got a more general upper bound on the number of rational points of arbitrary varieties (possibly non-equidimensional). In addition, he proved that (1) is sharp by providing examples of varieties reaching this bound.

6.7. New families of fast elliptic curves

B. Smith has pioneered the use of mod-p reductions of Q-curves to produce elliptic curves with efficient scalar multiplication algorithms—which translates into faster encryption, decryption, signing, and signature verification operations on these curves. A theoretical article was presented at ASIACRYPT 2013 [7], and a longer version was submitted (upon invitation) to the Journal of Cryptology. The theory was put into practice in collaboration with Craig Costello (Microsoft Research) and Huseyin Hisil (Yasar University). Their resulting publicly available implementation, which represents the state of the art in constant-time (side-channel conscious) elliptic curve scalar multiplication on 64-bit Intel platforms at the 128-bit security level, can carry out a constant-time scalar multiplication in 145k cycles on Ivy Bridge architectures. This work appeared in EUROCRYPT 2014 [21].

6.8. New results for solving the discrete logarithm problem

Recent results of R. Barbulescu, P. Gaudry, A. Joux, and E. Thomé seem to indicate that solving the discrete logarithm problem over finite fields of small characteristic is easier than was precedently thought. F. Morain and A. Guillevic, joined by R. Barbulescu and P. Gaudry, embarked on an attempt to assess the security of the discrete logarithm problem in a closely related context: that of finite fields with large characteristic and small degree. Improving on the methods of A. Joux, R. Lercier and others, they found new algorithms to select polynomials for the Number Field Sieve – the algorithm of choice in this setting. Moreover, a clever study of the algebraic properties of the fields used (e.g., algebraic units), enabled them to break the world record for the case of $GF(p^2)$, soon to be followed by new cases. This work is described in [31], and part of it is currently submitted.

6.9. Quantum Integer Factorization

Together with two researchers in quantum physics (F. Grosshans and T. Lawson), F. Morain and B. Smith have been working on the number theoretical postprocessing in Shor's algorithm. A preprint is being written.

GRAPHIK Project-Team

6. New Results

6.1. Highlights of the Year

- *Michael Thomazo was awarded the AFIA Prize 2014 (French Association for Artificial Intelligence) for his PhD entitled "Conjunctive Query Answering Under Existential Rules - Decidability, Complexity, and Algorithms" defended in October 2013. He was also awarded the first accessit of Gilles Kahn Prize 2014 by the SIF (French Society for Computer Science) [14].*
- *Madalina Croitoru and Alain Gutierrez were awarded the Best Technical Paper of SGAI-2014 for "On Ontological Expressivity and Modelling Argumentation Schemes using COGUI", in collaboration with Wael Hamdan, Rady Khazem and Ghaida Rebdawi .*
- *Abdallah Arioua was awarded the Best Student Paper Award of SGAI-2014 for "Query Failure Explanation in Inconsistent Knowledge Bases: A Dialogical Approach" in collaboration with Nouredine Tamani, Madalina Croitoru and Patrice Buche .*

BEST PAPERS AWARDS :
[36] **AI'2014: Thirty-fourth SGAI International Conference on Artificial Intelligence**. W. HAMDAN, R. KHAZEM, G. REBDAWI, M. CROITORU, A. GUTIERREZ, P. BUCHE.
[28] **AI'2014: 34th SGAI International Conference on Innovative Techniques and Applications of Artificial Intelligence**. A. ARIOUA, N. TAMANI, M. CROITORU, P. BUCHE.

6.2. Ontology-Based Query Answering with Existential Rules

Participants: Jean-François Baget, Fabien Garreau, Mélanie König, Michel Leclère, Marie-Laure Mugnier, Swan Rocher, Federico Ulliana.

Ontolology-based query answering (and more generally *Ontology-Based Data Access*, *OBDA*) is a new paradigm in data management, which takes into account inferences enabled by an ontology when querying data. In other words, the notion of a database is replaced by that of a knowledge base, composed of data (also called facts) and of an ontology. In this context, *existential rules* (also called Datalog+) have been proposed to represent the ontological component [59], [58]. This expressive formalism generalizes both description logics used in OBDA (such as $\mathcal{EL}$ and DL-Lite), which form the cores of so-called tractable profiles of the Semantic Web ontological language OWL2) and Datalog, the language of deductive databases. Since about five years, we have been studying the theoretical foundations of this framework (mainly concerning decidability and complexity) and developing associated algorithmic techniques. We have started the development of a platform dedicated to OBDA with existential rules (see section 5.2).

Before presenting this year' results, we recall the two classical ways of processing rules, namely forward chaining and backward chaining, also known as "materialization" and "query rewriting" in the OBDA setting. In forward chaining, the rules are applied to enrich the initial data and query answering can then be solved solved by evaluating the query against the "saturate" database (as in a classical database system *i.e.*, with forgetting the rules). The backward chaining process can be divided into two steps: first, the initial query is rewritten using the rules into a first-order query (typically a union of conjunctive queries, UCQ); then the rewritten query is evaluated against the initial database (again, as in a classical database system). Since entailment is not decidable with general existential rules, both forward and backwards processes may not halt.

6.2.1. Improvement of Query Rewriting Algorithms

These last two years, we designed and implemented a query rewriting algorithm that takes as input a set of existential rules and a UCQ q and outputs a UCQ, which is a sound and complete rewriting of q, whenever such a rewriting exists [60], [61], [62]. This year's main improvement to this algorithm is the implementation of a unifier able to process rules without decomposing their head into single atoms. This improvement appeared to be have a very high impact on the efficiency of query rewriting (up to 274 quicker on an ontology where 32% of the rules have a head composed of two atoms instead of a single one). Beside, much effort has been devoted to experiments: to find appropriate benchmarks, to build a translator from the Semantic Web format OWL/OWL2 to our existential rule format dlgp (since most existing ontologies are available in OWL/OWL2 format), to select existing tools to compare with, run them, finally compare tools on several criteria.

- *Results partially published in the Semantic Web Journal [22].*

Query rewriting techniques have the interest of being independent from the data. However, a main bottleneck is that the size of the rewritten query can be exponential in the size of the original query, hence the produced rewriting maybe not usable in practice. A well-known source of combinatorial explosion are some very simple rules, which form the core of any ontology, typically expressing concept and relation hierarchies, concept properties and relation signatures. We have proposed a rewriting technique, which consists in compiling these rules into a preorder on atoms and embedding this preorder into the rewriting process. This allows to compute compact rewritings that can be considered as "pivotal" representations, in the sense that they can be easily translated into different kinds of queries that can be evaluated by different kinds of database systems. The provided algorithm computes a sound, complete and minimal UCQ rewriting, if one exists. Experiments show that this technique leads to substantial gains in the query rewriting process, in terms of size and runtime, and scales on very large ontologies (several ten thousands of rules).

- *Results not published yet. Reported in Mélanie König's PhD thesis [17].*

6.2.2. A Better Approximation of Chase Termination for Existential Rules and their Extension to Non-monotonic Negation

Forward chaining with existential rules is known as the *chase* in databases. Various acyclicity notions ensuring chase termination have been proposed in the knowledge representation and databases. Acyclicity conditions found in the literature can be classified into two main families: the first one constrains the way existential variables are propagated during the chase and the second one constrains dependencies between rules *i.e.*, the fact that a rule may lead to trigger another rule. These conditions are based on different graphs, but all of them can be seen as forbidding "dangerous" cycles in the considered graph. We defined a new family of graphs that allows to unify and strictly generalize these acyclicity notions without increasing worst-case complexity.

Second, we considered the extension to existential rules with nonmonotonic negation under stable model semantics and further extended acyclicity results obtained in the positive case by exploiting negative information.

- *This work is part of Fabien Garreau and Swan Rocher's PhD theses. Results published at the European Conference on Artificial Intelligence (ECAI 2014)[30](long version as an arXiv report) and at the Workshop on Non-monotonic Reasoning (NMR 2014) [31].*

6.2.3. Detailed Results and Complements on Query Answering under Greedy Bounded-Treewidth Sets of Existential Rules

The family of greedy bounded-treewidth sets of existential rules (gbts) is an expressive class of rules for which entailment is decidable. This decidability property relies on a structural property of the saturation by the chase (*i.e.*, the set of inferred facts): for any initial set of facts, the saturation of these facts has a bounded treewidth (where the treewidth is computed on a graph associated with a set of atoms). Moreover, a tree decomposition of bounded width can be incrementally built during the chase. This family generalizes the important family of guarded existential rules, which itself generalizes Horn description logics used in OBDA.

In papers published at IJCAI 2011 and KR 2012, we studied the complexity of entailment under gbts rules as well as under known subclasses of gbts (with respect to data, combined and query complexity) and provided a generic algorithm with optimal worst-case complexity. This year, we finally completed a long report (75 pages) containing the detailed proofs of the results, some of them being very technical; in this report, we also clarified and reformulated the description of the generic algorithm, according to Michael Thomazo's PhD thesis (defended in October 2013); finally, we complemented the landscape of gbts classes by studying the complexity of all subclasses obtained by combining the syntactic criteria which define already known classes.

- *Results available as an arXiv report [56]. Submitted to a major journal in Artificial Intelligence. In collaboration with Sebastian Rudolph (TU Dresden) and Michael Thomazo (now postdoctoral student in Sebastian Rudolph's group).*

6.2.4. Extracting Bounded-level Modules from Deductive RDF Triplestores

The Semantic Web is consolidating a legacy of well-established knowledge bases spanning from life sciences, to geographic data and encyclopedical repositories. Today, reusing knowledge and data available online is vital to ensure a coherent development of the Semantic Web, thereby capitalizing on the efforts made in the last years by many institutions and domain experts to publish quality information.

In this paper we studied how to extract modules from RDF knowledge bases equipped with Datalog inference rules, we called Deductive RDF Triplestores. A module is a Deductive RDF Triplestore entailed from the reference system, which is defined upon a restricted vocabulary (or signature). We proposed a new semantics for bounded-level modules allowing to control their size, and then presented extraction algorithms compliant with the novel semantics. This feature is helpful since many ontologies are extremely large, while users often need to reuse only a small part of resources in their work.

This work was partially carried out before the arrival of Federico Ulliana at GraphIK. For the future, we plan to study module extraction for knowledge bases equipped with existential rules, which extend the rules considered here.

- *Results published at the Twenty-Ninth AAAI Conference on Artificial Intelligence (AAAI 15) [44]. In collaboration with Marie-Christine Rousset from LIG (University of Grenoble).*

6.2.5. Axiomatisation of Consistent Query Answering via Belief Revision

This work takes place in the OBQA setting where a query is being asked over a set of knowledge bases defined over a common ontology. When the union of knowledge bases along with the ontology is inconsistent, several semantics have been defined which are tolerant to inconsistency. These semantics all rely on computing repairs, *i.e.*, maximal (in terms of set inclusion) consistent subsets of the data set. They have been studied from a productivity point of view and a complexity point of view. We take a new point of view to define axiomatic characterisations of two such semantics, namely IAR (Intersection of All Repairs) and ICR ((Intersection of Closed Repairs). We argue that such characterisation can provide an alternative way of comparing the semantics and new insights into their properties. Furthermore such axiomatisation can be used when proposing a generalisation of inconsistency tolerant semantics. In order to provide the axiomatic characterisations we define belief revision operators that correspond to IAR and ICR.

- *Work published at [43]. In collaboration with Ricardo Rodriguez from University of Buenos Aires.*

6.3. Reasoning with Imperfect Information and Priorities

Participants: Abdallah Arioua, Patrice Buche, Madalina Croitoru, Jérôme Fortin, Souhila Kaci, Namrata Patel, Tjitze Rienstra, Nouredine Tamani, Rallou Thomopoulos.

This work focuses on two main notions, namely *argumentation systems*, which allow to represent and deal with conflicting information, and formalisms to represent *preferences*, which allow to compare possible outcomes in decision making and recommender systems.

6.3.1. Fundamental Aspects of Argumentation

A Dung-style argumentation framework aims at representing conflicts among elements called arguments. The basic ingredients of this framework is a set of arguments and a Boolean abstract (*i.e.*, its origin is not known) binary defeat relation on these arguments. This abstract framework can be instantiated in different ways, by representing arguments in a given knowledge representation formalism, which allows to take the semantics of arguments into account in the computation of the defeat relation.

Preference-based argumentation frameworks are instantiations of Dung's framework in which the defeat relation is derived from an attack relation and a preference relation over the arguments. Recently, Dung's framework has been extended in order to consider the strength of the defeat relation *i.e.*, to quantify the degree to which an argument defeats another argument. In this work, we instantiated this extended framework by a preference-based argumentation framework with a valued preference relation. As particular cases, the latter can be derived from a weight function over the arguments or a Boolean preference relation. We showed under some reasonable conditions that there are "less situations" in which a defense between arguments holds with a valued preference relation compared to a Boolean preference relation. Finally, we provided some conditions that the valued preference relation shall satisfy when it is derived from a weight function.

- *This is a joint work with Christophe Labreuche from Thales and published in [20]*

We also considered an extension to argumentative frameworks based on fuzzy set theory. The knowledge base is fuzzified to allow agents expressing their expertise (facts and rules) attached with grades of importance in the unit interval. Arguments are then attached with a strength score aggregating the importance expressed on their facts and rules. Extensions, corresponding to subsets of consistent arguments, are also attached with forces computed based on their strong arguments. The forces are used then to rank extensions from the strongest to the weakest one, upon which decisions can be made. We have also shown that the strength preference relation defined over arguments is reasonable according to classical rationality postulates and our fuzzy logical argumentation system can be seen as a computationally efficient instantiation of the generic model of structured argumentation framework. We furthered our theoretical research and demonstrate the added value of our approach in the practical setting of the European project EcoBioCap (see Sect.8.2).

- *Work published in IPMU 2014 [47] and Fuzz IEEE 2014 [46].*

One instantiation, among many others, of Dung's framework consists in constructing the arguments from a set of propositional logic formulas. Thus an argument is seen as a reason for or against the truth of a particular statement. Despite its advantages, the argumentation approach for inconsistency handling also has important shortcomings. More precisely, in some applications what one is interested in are not so much only the conclusions supported by the arguments but also the precise explanations of such conclusions. We showed that argumentation framework applied to classical logic formulas is not suitable to deal with this problem. On the other hand, intuitionistic logic appears to be a natural alternative candidate logic (instead of classical logic) to instantiate Dung's framework. We developed *constructive argumentation framework*. We showed that intuitionistic logic offers nice and desirable properties of the arguments. We also provided a characterization of the arguments in this setting in terms of minimal inconsistent subsets when intuitionistic logic is embedded in the modal logic S4.

- *This is a joint work with Yakoub Salhi from CRIL and published in [39]*

Lastly, we developed a model of abduction in abstract argumentation, where changes to an argumentation framework act as hypotheses to explain the support of an observation. We presented dialogical proof theories for the main decision problems (*i.e.*, finding hypotheses that explain skeptical/credulous support) and we showed that our model can be instantiated on the basis of abductive logic programs.

- *This work has been done in Tjitze Rienstra's thesis and published in [32].*

6.3.2. Use of Argumentation in Explanation, Querying and Decision Making

Besides work on the foundations of argumentation frameworks, we have studied the use of argumentation techniques in various tasks: explanation of query failure, reverse engineering, and decison making. These studies are mainly motivated by agri-food scenarii: bread conception, packaging conception, and durum wheat conception.

We have proposed an argumentation-based explanation for *query failure explanation* under the inconsistency tolerant semantics ICR in an Ontology-Based Data Access setting with existential rules. We used a rule-based language and we base our work on the equivalence between ICR-based query answering in inconsistent knowledge bases and sceptical acceptance of arguments. We proposed a multilevel explanation that exploits both the inference power of the logical language as well as arguments of dialectical nature. We also investigated an interactive argumentative approach where the process of explanation takes the form of a dialogue between the user and the reasoner.

- *Work published in COMMA 2014 [27] and SGAI 2014 [28] where it received the best student paper award.*

Within the framework of the European project EcoBioCap http://www.ecobiocap.eu about the design of next generation packagings using advanced composite structures based on constituents derived from the food industry, we have been developing a Decision Support System (DSS) for packaging material selection. [40], [49]. The DSS consists of two steps: (1) aggregating possibly conflicting needs expressed by several parties involved in the considered field and (2) querying a database of packagings with the resulting aggregation obtained at point (1). We instantiate for each need, called viewpoint or aspect, an argumentation system to reason about arguments solely expressed on it [45]. This will then be used to generate the query on the packaging database. To this aim we show how to instantiate ASPIC with the DLR-Lite logic modeling expert ontologies in this real world scenario [47].

- *Work published in AAMAS 2014 [45], IPMU 2014 [47], ICCS 2014 [40], and COMMA 2014 [49].*

Evaluating food quality is a complex process since it relies on numerous criteria historically grouped into four main types: nutritional, sensorial, practical and hygienic qualities. They may be completed by other emerging preoccupations such as the environmental impact, economic phenomena, etc. However, all these aspects of quality and their various components are not always compatible and their simultaneous improvement is a problem that sometimes has no obvious solution, which corresponds to a real issue for decision making. We propose a decision support method guided by the objectives defined for the end products of an agrifood chain. It is materialized by a backward chaining approach based on argumentation [47]. An extended version of this paper reporting on experimental results and expert evaluation has been published in Ecological Informatics [24].

- *Work published in IPMU 2014 [47], and Ecological Informatics 2014 [24].*

Knowledge elicitation, representation and reasoning explanation by / to non-computing experts has always been considered as a crafty task due to difficulty of expressing logical statements by non-logicians. We use the COGUI editor in order to elicit and represent argumentation schemes expressed using existantial rules within an inconsistent knowledge base. COGUI is a visual, graph based knowledge representation editor compatible with main Semantic Web languages (see Section 5.1). COGUI allows for default reasoning on top of ontologies. We investigate its use for modelling and reasoning using argumentation schemes and discuss the advantages of such representation. We show how this approach can be useful in the practical setting of EcoBioCap where the different argumentation schemes can be used to lead reasoning.

- *Work published in SGAI 2014 [36] where it received the best technical paper award. In collaboration with Wael Hamdan, Rady Khazem and Ghaisa Rebdawi from the Higher Institute of Applied Science and Technology (HIAST), Syria.*

6.3.3. Preferences

Qualitative and comparative preference statements of the form “prefer α to β” are useful components of many applications. This statement leads to the comparison of two sets of alternatives: the set of alternatives in which α is true and the set of alternatives in which β is true. Different ways are possible to compare two sets of objects leading to what is commonly known as preference semantics. The choice of the semantics to employ is important as they differently rank-order alternatives. Existing semantics are based on philosophical and non-monotonic reasoning grounds. In the meanwhile, they have been widely and mainly investigated by AI researchers from algorithmic point of view. We came to this problem from a new angle and completed existing theoretical investigations of the semantics. In particular, we provided a comparison of the semantics on the basis of their psychological plausibility by evaluating their closeness to human behavior.

- *This is a joint work with Eric Raufaste from CLLE and published in [38]*

There has been a growing interest in the study of preferences for their utility in solving problems related with decision making. Most of the preference representation languages developed in the literature are based on comparative preference statements since they offer a simple and intuitive way for expressing preferences. They can be further interpreted following different semantics, imparting a greater flexibility on how outcomes can be compared. So far the main objective has been to rank-order the set of outcomes given a set of comparative preference statements and one or several semantics. Tackling this problem from a different angle, we looked into the behavioral aspects of the preference semantics and statements by attempting to formalise the intuition behind them using postulates studied in preference logics and non-monotonic reasoning. We selected the postulates w.r.t. three criteria: coherence, syntax independence and inference. Thus, our analysis provided a means to determine those properties that are satisfied for a given preference semantics.

- *This work has been done in Namrata Patel's thesis and published in [21]*

Intelligent 'services' are increasingly used on e-commerce platforms to provide assistance to customers. Numerous preference elicitation methods developed in the literature are now employed for this purpose. However, it is commonly known that there is a real bottleneck in preference handling as concerns the elicitation of preferences because it does not cater to the wide range of preference representation languages available. Thus, as a first step in developing a decision-support tool using an AI based on such languages, this paper describes a preliminary study conducted to address this issue. We proposed a method of eliciting real-time user preferences expressed in natural language (NL) which can be formally represented using comparative preference statements complying with different semantics, and provided a proof of concept to demonstrate its feasibility. Since we developed NL resources to detect preference semantics, we also made a comparative study with existing resources to underline the peculiarities of our model.

- *This work has been done in Namrata Patel's thesis and published in [37]*

6.4. Semantic Data Integration

Participants: Michel Chein, Madalina Croitoru, Léa Guizol, Michel Leclère, Rallou Thomopoulos.

It often happens that different references (*i.e.*, data descriptions), possibly coming from heterogeneous data sources, concern the same real world entity. In such cases, it is necessary: (i) to detect whether different data descriptions really refer to the same real world entity and (ii) to fuse them into a unique representation. This issue has been been studied under various names: "record linking", "entity resolution", "reference resolution", "de-duplication", "object identification", "data reconciliation", etc., mostly in databases . It has become one of the major challenges in the Web of Data, where the objective is to link data published on the web and to process them as a single distributed database.

We investigate this problem in the specific context of bibliographic databases. Indeed, people working in bibliographical information systems have a lasting tradition of using norms and have integrated, along collections of documents notices (e.g. bibliographic records), collections of authority notices that categorize the different named entities used to describe documents (people, organizations, places, ...). In current databases, documents notices do not use directly the names of named entities to fill a particular field (author, editor, ...), but the unique identifier of the authority notice representing that named entity.

A few years ago, we began a collaboration with ABES (National Bibliographic Agency for Universities) to develop a method and a prototype to perform entity resolution between on one hand the authors of a new bibliographic record, and, on the other the authority references of an authority catalog (and namely the Sudoc catalog from the ABES agency). A problem with this approach is that it relies upon pre-established links between bibliographic records and authority notices. However, our experimentation and evaluation have shown that many existing links were erroneous, and thus led to the propagation of new linkage errors. We have thus began to work on methods and tools to repair linkage errors in bibliographical databases. The first step of our approach was to build a knowledge-base over an ontology (based on the international standards FRBR and CIDOC-CRM) aiming at representing bibliographic data (an RDFS base) as well as librarian knowledge.

From that, we developed a methodological framework allowing to design rules concluding on the coreference or the difference between entities of the bibliographic knowledge base. This framework was implemented in Cogui.

6.4.1. An Original Methodology to Compute Coreference and Difference Links

Our methodology can be briefly summarized as follows. The first step consists in computing "sure" links. In the second step, authority notices are enriched by information that comes from bibliographic notices to which they are linked by sure links. In the third step, Datalog rules that conclude on coreference or difference are triggered. The results are used to compute new sure links. These steps are iterated until stability *i.e.*, no new sure link is discovered. More specifically, the Datalog rules are the following form. The body of a rule is a conjunction of similarity criteria on attributes and its head states the coreference or the difference of two individual entities with a specific confidence level (represented as a symbolic value). We are currently instantiating this methodology for the Sudoc catalog, jointly with the ABES librarians, which will allow them to evaluate it.

6.4.2. Partioning Semantics for Link Discovery in Bibliographic Knowledge Bases

With the aim of evaluating and improving the quality of links in bibliographical knowledge bases, we have developed a decision support system based on partitioning semantics. The novelty of our approach consists in using symbolic values criteria for partitioning and suitable partitioning semantics. We have investigated the limits of those partitioning semantics: how the characteristics of the input (objects and criteria) influences characteristics of the result, namely correctness of the result and execution time. We have also evaluated and compared the above mentioned semantics on a real qualitative sample. This sample is issued from the catalogue of French university libraries (SUDOC) maintained by ABES.

- *This work is part of Lea Guizol's PhD thesis [16]. Work published in Fuzz IEEE 2014 [46].*

6.4.3. Key Discovery on the Semantic Web

Many techniques were recently proposed to automate the linkage of RDF datasets. Predicate selection is the step of the linkage process that consists in selecting the smallest set of relevant predicates needed to enable instance comparison. We call keys this set of predicates that is analogous to the notion of keys in relational databases. We have formally explained the different assumptions behind two existing key semantics (IC), and have evaluated experimentally these keys semantics by studying how discovered keys could help dataset interlinking or cleaning.

- *Work published in IC 2014 [50] and ICCS 2014 [29] in collaboration with Manuel Atencia and Jerome David from LIG, and Nathalie Pernelle, Fatiha Sais and Danai Symeonidou from LRI. See also the reconciliation-based approach in[23].*

6.4.4. Fusion of Linked Data

The problem of data fusion starts from reconciled datasets, whose objects are linked with semantic sameAs relations, as described above. We attempt to merge the often conflicting information of these reconciled objects in order to obtain unified representations that only contain the best quality information. We are studying an approach to determine the most appropriate value(s). Our method combines different quality criteria based on the value and its data source, and exploits, whenever possible, the ontology semantics, constraints and relations. Moreover we create a mechanism to provide explanations about the quality of each value, as estimated by our system. To achieve this, we generate annotations used for traceability and explanation purposes.

- *Work described in the Qualinca deliverable 4.2 research report, and accepted for publication in EGC'2015 : "Linked Data Annotation and Fusion driven by Data Quality Evaluation" (authors: Ioanna Giannopoulou, Fatiha Saïs from LRI, and Rallou Thomopoulos)*

.

HEPHAISTOS Team

6. New Results

6.1. Highlights of the Year

Yves Papegay received a "Wolfram Innovator Award" in December 2014

6.2. Robotics

6.2.1. Cable-driven parallel robots (CDPR)

6.2.1.1. Analysis of Cable-driven parallel robots

Participants: Alessandro Berti, Laurent Blanchet, Houssein Lamine, Jean-Pierre Merlet [correspondant], Yves Papegay, Rémy Ramadour.

We have continued the analysis of suspended CDPRs for control and design purposes. For control it is essential to determine the current pose of the robot for given leg lengths. This forward kinematic problem (FK) is usually very complex and admits several solutions. For parallel robot with rigid legs we have established the important property (P) that the FK may be solved in real-time i.e. being given the leg lengths ρ and platform pose $\mathbf{X}$ it is possible to determine the single pose $\mathbf{X_1}$ that can be reached from $\mathbf{X}$ if the leg lengths has been changed to $\rho + \Delta\rho$ provided that $\Delta\rho$ satisfies some properties. For CDPR with sagging cables determining all the FK solutions is more complex but we have proposed the first algorithm to solve it the for a full scale model of sagging cables [24]. For CDPR with non sagging cables the problem is also very complex because we cannot make any assumption on the number of cables under tension i.e for a CDPR with m cables we have to solve all the FK problems for all possible set of cables under tension from 1 to m and as soon as this number is lower than 6 the system of equations is much larger than for classical robots. We have however been able to propose an interval analysis based algorithm that allow one to get all the solutions [18]. But we have also shown that for non sagging cables the property (P) does not hold. Indeed it requires that the system of equations that governs the FK remains the same at all time. But for CDPR this system depend on the set of cables under tension (which is called the *cable configuration CC*) and it may change when the cable lengths change from 'ρ to $\rho + \Delta\rho$, even for redundant CDPR [23]. If the CC changes at some point the pose solution of the FK together with the cable tensions will differ from the one that is obtained when assuming no change in the CC. This has a drastic effect on control as we have now a system whose state equations may change over time but also on design as in the new CC the cable tensions may be quite different from the expected one. Hence property (P) will hold if and only if we are able to show that there will not be any change in the CC during the change of the cable lengths and therefore it is crucial to detect CC changes. But this require to fully simulate the discrete-time control laws together with the behavior of the coiling system. We have been able to implement a simulation tool that tracks a trajectory for the robot for arbitrary control laws and coiling system model [22], [25]. The principle of the algorithm is to determine if on a time interval $[t, t + \delta t]$ the solution of the FK with the current CC satisfies (P) by using Kantorovitch theorem. If this is not the case Δt is divided by 2 and the process is repeated. We then check if there is a time t_1 in $[t, t + \delta t]$ for which the tension of a cable in the CC may become 0 If there is no such t_1 for any cable in the current CC, then it will be the CC at time $t + \delta t$ and we may compute the pose and cable tensions at any time in $[t, t + \delta t]$. If there is a least one such t_1 (and there may be several $t_1^i, t_1^j, \ldots$ as we consider each cable in the CC) we order these times by increasing values and check sequentially if a cable tension become negative with the current CC at time $(t_1^i + t_1^{i+1})/2$. If yes we determine what can the CC at this time by looking at all possible CC. As soon as the new CC at time t_1^i has been determinated the simulation can go on. Implementing this algorithm has been difficult mainly for numerical reasons: the accuracy of the calculation may sometime exceed the floating point accuracy and we have to resort to symbolic computation and extended arithmetics. Our tests have shown that indeed CC changes may occur on trajectories: on a typical trajectory up to 10 different CC will appear with 5 or 6 cables under tension. These results have been confirmed experimentally on a prototype at LIRMM.

6.2.1.2. Tool for Agencement Analysis and Synthesis of CDPRs

Participants: Laurent Blanchet, Jean-Pierre Merlet [correspondant].

HEPHAISTOS has been working on tools to design the layout and geometry of CDPRs, while accounting for numerical errors as well as practical errors – actual position of the winches, of the attachments on the platform, errors of the controllers, of the cables, etc. Within this work, collision analysis plays an important role. Indeed the concept of cable robot aims to increase the workspace that is restricted for robots having rigid legs but interferences may reduce this workspace. Two types of interference analysis approaches exist: intersection of numerically-mapped boundaries (InB) and distance between features (DbF). The two sets of interference types that can be analysed using these approaches are distinct but overlapping. The first approach greatly benefits from Inria's computational geometry research and particularly from the AABB tree algorithms implemented in CGAL. Algorithms and implementation based on those were developed, along with several new algorithm and implementation to extend the scope of intersection types, and thus, of interference types. Algorithms to improve efficiency of given intersection types were also developed. We have already used the second approach, DbF, to develop algorithms for leg interference of parallel robots that are very efficient for non deformable cables but now well adapted for sagging cables. An interference detection algorithm has been developed and implemented for a restricted scope of applications [10], and research is on-going for a more generic case.

6.2.1.3. Visual-servoing of a parallel cable-driven robot

Participants: Rémy Ramadour, Jean-Pierre Merlet [correspondant], François Chaumette [correspondant].

The last two years, we studied how visual servoing could improve accuracy, controllability and performance of cable-driven parallel robots [13]. Previous works on this domain showed very interesting results but some issues remained to be investigated , such as :

- ratio accuracy/workspace : cable-driven parallel robots are known to allow a large reachable workspace, but also to have complex geometric and dynamic models which affect the accuracy. Using visual-servoing in a closed-loop scheme, we were able to enhance the accuracy by a factor of ten, allowing to manipulate daily-life objects in a whole living room.
- image-based joint-space control : in order to reach a desired pose, the usual method involves several computing and evaluations of both the Jacobian matrix of the manipulator and the interaction matrix linking visual features to the displacements of the end-effector. We designed a control scheme, based on an iterative updating using the Broyden update law, in order to link the visual features directly to the joint coordinates. This scheme is less sensitive to model uncertainties and require much less computing.
- stability of the command law : classical control laws ignore cable configuration effects that change the pose of the platform. We have proposed a counter-intuitive strategy: the robot `MARIONET-ASSIST` we are using has a specific geometry that allow to predict which cables set may be under tension for a given trajectory i.e. we are able to split the trajectory in parts for which we know all possible cables configurations. Among them we select the one that optimize an accuracy criteria and we enforce it by forcing the cables not part of the configuration to be slack by adding a sufficient amount of length to their nominal values. It allowed to enhance both the stability and the accuracy of a vision-based control scheme [26].

We also used interval analysis in order to guaranty every step of the process, in order to provide safety and reliability of our methods, as the robots that we use were initially deployed in the context of assistive technologies.

Finally, simulations and experiments on prototypes were conducted and presented in order to validate the mentioned results. However, the prototype that we used presents a very particular configuration (all wires are connected to the same point on the end-effector, allowing only translational movements), further works may be required in order to test our methods for a wider variety of cable-driven parallel robots.

6.2.1.4. Cable-Driven Parallel Robots for additive manufacturing in architecture

Participant: Yves Papegay.

Easy to deploy and to reconfigure, dynamically efficient in large workspaces even with payloads, cable-driven parallel robots are very attractive for solving deplacement and positioning problems in architectural building at scale 1 and seems to be a good alternative to crane and industrial manipulators in this area.

In a collaboration with CNAM and Ecole Nationale Supérieure d'Architecture Paris-Malaquais, we aim to design and realize a CDPR of large size as a proof of concept in additive manufacturing of building based on ultra-high performance concrete.

Challenges are modeling and control to get enough accuracy.

6.2.2. Assistance robotics

This is now the core of our activity and our work on CDPR is deeply connected to this field as they are an efficient solution for mobility assistance, a high priority for the elderly, helpers and medical community. We have presented our vision of assistance robotics in several occasions [20], [21], [19].

6.2.2.1. Assessment of elderly frailty

Participants: Karim Bakal, Jean-Pierre Merlet.

The assessment of elderly frailty is a difficult concept because it involves the physical capacities of a person and its environment (health-care services, families, funds...). We consider the assessment of upper limb capabilities by looking at the joint torques τ of the arm and the maximal force F that can be exerted by the hand, which are related by the equation

$$\tau = \mathbf{J}^{\mathbf{T}} F \tag{7}$$

where $\mathbf{J}$ is a matrix which depends only upon the configuration of the arm. This equations constitutes an underconstrained linear system. In biomechanics the torque τ is measured together with the configuration of the arm and the force F is evaluated by using the method of Chiacchio, that involves the pseudo-inverse of $\mathbf{J}^{\mathbf{T}}$ to calculate F. But there are several uncertainties that are neglected when using this method: the measurement errors on τ and on the configuration of the arm together with uncertainties on the physical parameters of the arm (such as the length of the bones). The method of Chiacchio provides one of the possible solutions of equation (2) and not necessary the one corresponding to the force at the hand. We use another approach based on interval analysis. We assume that all uncertainties may be bounded (τ is an interval vector τ_m, $\mathbf{J}^{\mathbf{T}}$ is an interval matrix) so that equation (2) become an interval linear system. Interval analysis then allows one to determine an approximation as accurate as wanted of the set F_s of all forces F that satisfy the equation and therefore this set includes the real force at the hand. Now assume that with the same arm configuration we measure the force at the hand, here again with some bounded uncertainties (i.e. F is an interval vector F_m). Here again we may use interval analysis applied on equation (2) in order to determine an interval vector τ_v for the τ that is guaranteed to include the real τ. Furthermore τ must be included in the intersection τ_i of τ_v and τ_m while F must be included in the intersection F_i of F_m and F_s. If τ_i is strictly included in τ_m, then we may compute a better approximation of F_s. Reciprocally if F_i is strictly included in F_m we will get a better τ_v. If one of these situation occurs we repeat the process until no significant improvement of F_s or τ_v is obtained. In a second step we consider that the uncertainties that lead to uncertainties in the matrix $\mathbf{J}^{\mathbf{T}}$ are constrained as we have to satisfy $\tau_v = \mathbf{J}^{\mathbf{T}} F_s$. Here again we use interval analysis to determine if this constraint does not allow to reduce the size of the interval on the physical parameters in which case we may obtain a new $\mathbf{J}^{\mathbf{T}}$ that is included in the initial one. In turn this may allow to obtain better τ_v and F_s. The process stops when no improvement has been obtained for F_s, τ_v and the physical parameters.

To test this approach the right upper limb joint torque of 10 males and the force capacity at the right hand was measured by a dynamometer (Biodex III, Biodex Medical Systems) and respectively by a 6-axis load sensor during an experiment performed at HandiBio laboratory. The configuration of the upper limb was measured with a motion capture system (Qualisys, Sweden). The approach is currently being evaluated.

6.2.2.2. Walking analysis

Participants: Claire Maillard, Ting Wang, Jean-Pierre Merlet [correspondant].

The walkers of the ANG family allow one to determine accurately the trajectory of the walker and therefore to analyse the walking of the user. We have used this property for performing until mid 2013 a large scale experiment: 23 young adults and 25 elderly people (> 69 years) were asked to walk along with two reference trajectories with the help of the walker. The objective of this research is to develop walking quality index and examine if the walker may be used to monitor the health state of elderly people at home. We compared and statistically analyzed the walking patterns of the two groups of people. The results show that it is possible to obtain new indicators by using the walker measurements [9],[14]. Next step will be to perform a similar analysis for a sit-to-stand (STS) exercice and to test our approach in two rehabilitation centers, MATIA in Spain (in the framework of the RAPP project) and Centre Héliomarin de Vallauris.

A start-up plan was proposed in November 2014 to transfer the walking analysis technology of Hephaistos with the ANG walker. In order to study the feasibility of our plan, we have interviewed Patrick Nenert (Kiné, Centre Hélio-Marin), Françoise Dubourgeois (DR, EPHAD) and Sophie Morgenstern (Métropole NCA, Living Lab Paillon 2020) about their impression of the walker and the possibility of the future collaboration with them. Several contact with local actors of the silver economy sector have already been established : Livinglab Paillon2020 (Nice), CIU-santé, as well as with research lab for collaboration on future projects (Lapcos, I3M, Gredeg).

6.2.2.3. Design and evaluation of assistive devices, ethics

Participants: Marc Beninati, Bernard Senach [correspondant], Jean-Pierre Merlet.

Providing appropriate support, services and information to the elderly, to their caregivers and to the medical profession, through a fleet of communicating devices must rely on a structured processes. A generic design and evaluation framework is being elaborated and will be validated through field experiments.

Assistance robotics raises many ethical questions. We started reflection about conducting experiments with frail and old people. A listing of questions to be addressed at each step of an experiment has been written (internal document). We have also hired a joint PhD student with University Bologna about the legal aspects of assistance robotics and we plan to organize a national forum on this topic with Nathalie Nevejans from University of Douai.

6.3. Miscellaneous results

6.3.1. Symbolic tools for modeling and simulation

Participant: Yves Papegay.

This activity is the main part of a long-term ongoing collaboration with Airbus whose goal is to directly translate the conceptual work of aeronautics engineers into digital simulators to accelerate aircraft design.

An extensive modeling and simulation platform has been designed which includes a dedicated modeling language for the description of aircraft dynamics models in term of formulae and algorithms, and a symbolic compiler producing as target an efficient numerical simulation code ready to be plugged into a flight simulator, as well as a formatted documentation compliant with industrial requirements of corporate memory.

Technology demonstrated by our prototype has been transferred to our industrial partner in 2012 when final version of our modeling and simulation environment has been delivered to Airbus in November 2012. Developer level know-how has been transferred in 2013 to a software company in charge of industrialization and maintenance of the modeling and simulation environment.

In 2014, we have worked again on several enhancements and extension of functionalities, namely to enhance the performances and the numerical quality of the generated C simulation code, and ease the integration of our environment into the Airbus toolbox.

HIEPACS Project-Team

6. New Results

6.1. Highlights of the Year

In the context of HPC-PME initiative, we started a collaboration with ALGO'TECH INFORMATIQUE and we have organised one of the first PhD-consultant action implemented by Xavier Lacoste led by Pierre Ramet. ALGO'TECH is one of the most innovative SMEs (small and medium sized enterprises) in the field of cabling embedded systems, and more broadly, automatic devices. The main target of the project is to validate the possibility to use the sparse linear solvers of our team in the area of electromagnetic simulation tools developped by ALGO'TECH. This collaboration will be developed next year in the context of the European project FORSTISSIMO. The principal objective of FORTISSIMO is to enable European manufacturing, particularly SMEs, to benefit from the efficiency and competitive advantage inherent in the use of simulation.

As a conclusion of the OPTIDIS project we organized the first International Workshop on Dislocation Dynamics Simulations that was devoted to the latest developments realized worldwide in the field of Discrete Dislocation Dynamics simulations. This international event held in December 10th to the 12th at "Maison de la Simulation" in Saclay, France and attracted 55 participants from many different countries including England, Germany, France, USA, ... The workshop gathered most of the active researchers working on dislocation dynamics from numerical simulations to experimentatios. Thanks to the success of this workshop, a second one will be scheduled in England during 2016.

6.2. High-performance computing on next generation architectures

6.2.1. *Composing multiple StarPU applications over heterogeneous machines: a supervised approach*

Enabling HPC applications to perform efficiently when invoking multiple parallel libraries simultaneously is a great challenge. Even if a uniform runtime system is used underneath, scheduling tasks or threads coming from different libraries over the same set of hardware resources introduces many issues, such as resource oversubscription, undesirable cache flushes or memory bus contention.

This work presents an extension of `StarPU`, a runtime system specifically designed for heterogeneous architectures, that allows multiple parallel codes to run concurrently with minimal interference. Such parallel codes run within *scheduling contexts* that provide confined execution environments which can be used to partition computing resources. Scheduling contexts can be dynamically resized to optimize the allocation of computing resources among concurrently running libraries. We introduce a *hypervisor* that automatically expands or shrinks contexts using feedback from the runtime system (e.g. resource utilization). We demonstrate the relevance of our approach using benchmarks invoking multiple high performance linear algebra kernels simultaneously on top of heterogeneous multicore machines. We show that our mechanism can dramatically improve the overall application run time (-34%), most notably by reducing the average cache miss ratio (-50%).
This work is developed in the framework of Andra Hugo's PhD. These contributions have been published in the international journal of High Performance Computing Applications [21].

6.2.2. *A task-based $\mathcal{H}$-Matrix solver for acoustic and electromagnetic problems on multicore architectures*

$\mathcal{H}$-Matrix is a hierarchical, data-sparse approximate representation of matrices that allows the fast approximate computation of matrix products, LU and LDL^T decompositions, inversion and more. This representation is suitable for the direct solution of large dense linear systems arising from the Boundary Element Method in $O(N \log_2^{\alpha}(N))$ operations. This kind of formulation is widely used in the industry for the numerical simulation of acoustics and electromagnetism scattering by large objetcs. Applications of this approach include aircraft noise reduction and antenna sitting at Airbus Group. The recursive and irregular nature of these $\mathcal{H}$-Matrix algorithms makes an efficient parallel implementation very challenging, especially when relying on a "Bulk Synchronous Parallel" paradigm. We have considered an alternative parallelization for multicore architectures using a task-based approach on top of a runtime system, namely StarPU. We have showed that our method leads to a highly efficient, fully pipelined computation on large real-world industrial test cases provided by Airbus Group.

This research activity has been conduced in the framework of the EADS-ASTRIUM, Inria, Conseil Régional initiative in collaboration with the RUNTIME Inria project, and is part of Benoit Lize's PhD.

6.2.3. A task-based 3D geophysics application

Reverse Time Migration (RTM) technique produces underground images using wave propagation. A discretization based on the Discontinuous Galerkin (DG) method unleashes a massively parallel elastodynamics simulation, an interesting feature for current and future architectures. We have designed a task-based version of this scheme in order to enable the use of manycore architectures. At this stage, we have demonstrated the efficiency of the approach on homogeneous and cache coherent Non Uniform Memory Access (ccNUMA) multicore platforms (up to 160 cores) and designed a prototype version of a distributed memory version that can exploit multiple instances of such architectures. This work has been conducted in the context of the DIP Inria-Total strategic action in collaboration with the MAGIQUE3D Inria project and thanks to the long-term visit of George Bosilca funded by TOTAL. Geroge's expertise ensured an optimum usage of the PaRSEC runtime system onto which our task-based scheme has been ported.

This work was presented during HPCC conference [27] as well as during a TOTAL scientific event [26].

6.2.4. Resiliency in numerical simulations

For the solution of systems of linear equations, various recovery-restart strategies have been investigated in the framework of Krylov subspace methods to address the situations of core failures. The basic underlying idea is to recover fault entries of the iterate via interpolation from existing values available on neighbor cores. In that resilience framework, we have extended the recovey-restart ideas to the solution of linear eigenvalue problems. Contrary to the linear system case, not only the current iterate can be interpolated but also part of the subspace where candidate eigenpairs are searched.

This work is developed in the framework of Mawussi Zounon's PhD funded by the ANR RESCUE. These contributions have been presented in particuler at the international SIAM workshop on Exascale Applied Mathematics Challenges and Opportunities [40] in Chicago and the Householder symposium [41] in Spa. Notice that theses activities are also part of our contribution to the G8 ESC (Enabling Climate Simulation at extreme scale).

6.2.5. Hierarchical DAG scheduling for hybrid distributed systems

Accelerator-enhanced computing platforms have drawn a lot of attention due to their massive peak computational capacity. Despite significant advances in the programming interfaces to such hybrid architectures, traditional programming paradigms struggle mapping the resulting multi-dimensional heterogeneity and the expression of algorithm parallelism, resulting in sub-optimal effective performance. Task-based programming paradigms have the capability to alleviate some of the programming challenges on distributed hybrid manycore architectures. In this work we take this concept a step further by showing that the potential of task-based programming paradigms can be greatly increased with minimal modification of the underlying runtime combined with the right algorithmic changes. We propose two novel recursive algorithmic variants for one-sided factorizations and describe the changes to the PaRSEC task-scheduling runtime to build a framework where the task granularity is dynamically adjusted to adapt the degree of available parallelism and kernel efficiency according to runtime conditions. Based on an extensive set of results we show that, with one-sided factorizations, i.e. Cholesky and QR, a carefully written algorithm, supported by an adaptive tasks-based runtime, is capable of reaching a degree of performance and scalability never achieved before in distributed hybrid environments.

These contributions will be presented at the international conference IPDPS 2015 [36] in Hyderabad.

6.3. High performance solvers for large linear algebra problems

6.3.1. Parallel sparse direct solver on runtime systems

The ongoing hardware evolution exhibits an escalation in the number, as well as in the heterogeneity, of the computing resources. The pressure to maintain reasonable levels of performance and portability, forces the application developers to leave the traditional programming paradigms and explore alternative solutions. PaStiX is a parallel sparse direct solver, based on a dynamic scheduler for modern hierarchical architectures. In this paper, we study the replacement of the highly specialized internal scheduler in PaStiX by two generic runtime frameworks: PaRSEC and StarPU. The tasks graph of the factorization step is made available to the two runtimes, providing them with the opportunity to optimize it in order to maximize the algorithm efficiency for a predefined execution environment. A comparative study of the performance of the PaStiX solver with the three schedulers - native PaStiX, StarPU and PaRSEC schedulers - on different execution contexts is performed. The analysis highlights the similarities from a performance point of view between the different execution supports. These results demonstrate that these generic DAG-based runtimes provide a uniform and portable programming interface across heterogeneous environments, and are, therefore, a sustainable solution for hybrid environments.

This work has been developed in the framework of Xavier Lacoste's PhD funded by the ANR ANEMOS. These contributions have been presented at the Heterogeneous Computing Workshop held jointly with the international conference IPDPS 2014 [32]. Xavier Lacoste will defend his PhD in February 2015.

6.3.2. Hybrid parallel implementation of hybrid solvers

In the framework of the hybrid direct/iterative MaPHyS solver, we have designed and implemented an hybrid MPI-thread variant. More precisely, the implementation relies on the multi-threaded MKL library for all the dense linear algebra calculations and the multi-threaded version of PaStiX. Among the technical difficulties, one was to make sure that the two multi-threaded libraries do not interfere with each other. The resulting software prototype is currently experimented to study its new capability to get flexibility and trade-off between the parallel and numerical efficiency. Parallel experiments have been conducted on the Plafrim plateform as well as on a large scale machine located at the USA DOE NERSC, which has a large number of CPU cores per socket.
This work is developed in the framework of the PhD thesis of Stojce Nakov funded by TOTAL.

6.3.3. Designing LU-QR hybrid solvers for performance and stability

New hybrid LU-QR algorithms for solving dense linear systems of the form $Ax = b$ have been introduced. Throughout a matrix factorization, these algorithms dynamically alternate LU with local pivoting and QR elimination steps, based upon some robustness criterion. LU elimination steps can be very efficiently parallelized, and are twice as cheap in terms of flops, as QR steps. However, LU steps are not necessarily stable, while QR steps are always stable. The hybrid algorithms execute a QR step when a robustness criterion detects some risk for instability, and they execute an LU step otherwise. Ideally, the choice between LU and QR steps must have a small computational overhead and must provide a satisfactory level of stability with as few QR steps as possible. In this work, we introduce several robustness criteria and we establish upper bounds on the growth factor of the norm of the updated matrix incurred by each of these criteria. In addition, we describe the implementation of the hybrid algorithms through an extension of the PaRSEC software to allow for dynamic choices during execution. Finally, we analyze both stability and performance results compared to state-of-the-art linear solvers on parallel distributed multicore platforms.
These contributions have been presented at the international conference IPDPS 2014 [30] in Phoenix. An extended version has been submitted to JPDC journal.

6.3.4. Divide and conquer symmetric tridiagonal eigensolver for multicore architectures

Computing eigenpairs of a symmetric matrix is a problem arising in many industrial applications, including quantum physics and finite-elements computation for automobiles. A classical approach is to reduce the matrix to tridiagonal form before computing eigenpairs of the tridiagonal matrix. Then, a back-transformation allows one to obtain the final solution. Parallelism issues of the reduction stage have already been tackled in different shared-memory libraries. In this work, we focus on solving the tridiagonal eigenproblem, and we describe a novel implementation of the Divide and Conquer algorithm. The algorithm is expressed as a sequential task-flow, scheduled in an out-of-order fashion by a dynamic runtime which allows the programmer to play with tasks granularity. The resulting implementation is between two and five times faster than the equivalent routine from the INTEL MKL library, and outperforms the best MRRR implementation for many matrices.
These contributions will be presented at the international conference IPDPS 2015 [34] in Hyderabad.

6.4. High performance Fast Multipole Method for N-body problems

Last year we have worked primarily on developing an efficient fast multipole method for heterogeneous architecture. Some of the accomplishments for this year include:

1. implementation of some new features in the FMM library ScalFMM: adaptive variants of the Chebyshev and Lagrange interpolation based FMM kernels, multiple right-hand sides, generic tensorial nearfield...
2. The parallelization and the FMM core parts rely on ScalFMM (OpenMP/MPI) which has been updated all year round. Finally, ScalFMM offers two new shared memory parallelization strategies using OpenMP 4 and `StarPU`.

6.4.1. Low rank approximations of matrices

New fast algorithms for the computation of low rank approximations of matrices were implemented in a -soon to be- open-source C++ library. These algorithms are based on randomized techniques combined with standard matrix decompositions (such as QR, Cholesky and SVD). The main contribution of this work is that we make use of ScalFMM parallel library in order to power the large amount of matrix to vector products involved in the algorithms. Applications to the fast generation of Gaussian random fields were adressed. Our methods compare good with the existing ones based on Cholesky or FFT and potentially outpass their performances for specific distributions. We are currently in the process of writing a paper on that topic. Extensions to fast Kalman filtering is now considered. This work is done in collaboration with Eric Darve (Stanford, Mechanical Engineering) in the context of the associate team FastLA.

6.4.2. Time-domain boundary element method

The Time-domain Boundary Element Method (TD-BEM) has not been widely studied but represents an interesting alternative to its frequency counterpart. Usually based on inefficient Sparse Matrix Vector-product (SpMV), we investigate other approaches in order to increase the sequential flop-rate. We present a novel approach based on the re-ordering of the interaction matrices in slices. We end up with a custom multi-vectors/vector product operation and compute it using SIMD intrinsic functions. We take advantage of the new order of the computation to parallelize in shared and distributed memory. We demonstrate the performance of our system by studying the sequential Flop-rate and the parallel scalability, and provide results based on an industrial test-case with up to 32 nodes [43], [28]. From the middle of year 2014, we started working on the TD FMM for the BEM problem. A non optimized version is able to solve the TD BEM with the FMM on parallel distributed nodes. All the implementations should be in high quality in the Software Engineering sense since the resulting library is going to be used by industrial applications.
This work is developed in the framework of Bérenger Bramas's PhD and contributes to the EADS-ASTRIUM, Inria, Conseil Régional initiative.

6.5. Efficient algorithmic for load balancing and code coupling in complex simulations

6.5.1. Dynamic load balancing for massively parallel coupled codes

In the field of scientific computing, load balancing is a major issue that determines the performance of parallel applications. Nowadays, simulations of real-life problems are becoming more and more complex, involving numerous coupled codes, representing different models. In this context, reaching high performance can be a great challenge. In the PhD of Maria Predari (started in october 2013), we develop new graph partitioning techniques, called co-partitioning, that address the problem of load balancing for two coupled codes: the key idea is to perform a "coupling-aware" partitioning, instead of partitioning these codes independently, as it is usually done. More precisely, we propose to enrich the classic graph model with *interedges*, that represent the coupled code interactions. We describe two new algorithms, called AWARE and PROJREPART, and compare them to the currently used approach (called NAIVE). In recent experimental results, we notice that both AWARE and PROJREPART algorithms succeed to balance the computational load in the coupling phase and in some cases they succeed to reduce the coupling communications costs. Surprisingly we notice that our algorithms do not degrade the global graph edgecut, despite the additional constraints that they impose. In future work, we aim at validating our results on real-life cases in the field of aeronautic propulsion. In order to achieve that, we plan to integrate our algorithms within the `Scotch` framework. Finally, our algorithms should be implemented in parallel and should be extended in order to manage more complex applications with more than two interacting models.

6.5.2. Graph partitioning for hybrid solvers

Nested Dissection has been introduced by A. George and is a very popular heuristic for sparse matrix ordering before numerical factorization. It allows to maximize the number of parallel tasks, while reducing the fill-in and the operation count. The basic standard idea is to build a "small separator" S of the graph associated with the matrix in order to split the remaining vertices in two parts P_0 and P_1 of "almost equal size". The vertices of the separator S are ordered with the largest indices, and then the same method is applied recursively on the two sub-graphs induced by P_0 and P_1. At the end, if k levels of recursion are done, we get 2^k sets of independent vertices separated from each other by $2^k - 1$ separators. However, if we examine precisely the complexity analysis for the estimation of asymptotic bounds for fill-in or operation count when using Nested Dissection ordering, we can notice that the size of the halo of the separated sub-graphs (set of external vertices belonging to an old separator and previously ordered) plays a crucial role in the asymptotic behavior achieved. In the perfect case, we need halo vertices to be balanced among parts. Considering now hybrid methods mixing both direct and iterative solvers such as `HIPS`, `MaPHyS`, obtaining a domain decomposition leading to a good balancing of both the size of domain interiors and the Scalable numerical schemes for scientific applications size of interfaces is a key point for load balancing and efficiency in a parallel context. This leads to the same issue: balancing the halo vertices to get balanced interfaces. For this purpose, we revisit the algorithm introduced by Lipton, Rose and Tarjan which performed the recursion of nested dissection in a different manner: at each level, we apply recursively the method to the sub-graphs But, for each sub-graph, we keep track of halo vertices. We have implemented that in the Scotch framework, and have studied its main algorithm to build a separator, called greedy graph growing.
This work is developed in the framework of Astrid Casadei's PhD. These contributions have been presented at the international conference HIPC 2014 [29] in Goa.

6.6. Application Domains

6.6.1. Dislocation dynamics simulations in material physics

6.6.1.1. Long range interaction

Various optimizations have been performed in the Dislocation Dynamics code OptiDis for the long-ranged isotropic elastic force and energy models using a Fast Fourier based Fast Multipole Method (also known as

Uniform FMM). Furthermore the anisotropic elastic force model was implemented using spherical harmonics expansions of angular functions known as Stroh matrices. Optimizations with respect to the crystallographic symmetries were also considered. Once the corresponding semi-analytic formulae for the force field are derived this method should compare well with existing approaches based on expanding the anisotropic elastic Green's function.

6.6.1.2. Parallel dislocation dynamics simulation

This year we have focused on the improvements of our hybrid MPI-OpenMP parallelism of the OptiDis code. More precisely, we have continued the development of the cache-conscious data structure to manage efficiently large set of data (segments and nodes) during all the steps of the algorithm. Moreover, we have tuned and improved our hybrid MPI-OpenMP parallelism to run simulations with large number of radiation induced defects forming our dislocation network. To obtain a good scalability, we have introduced a better load balancing at thread level as well as process level. By combining efficient data structure and hybrid parallelism we obtained a speedup of 112 on 160 cores for a simulation of half a million of segments.

These contributions have been presented in minisymposia at the 11th World Congress on Computational Mechanics [47], 7th MMM International Conference on Multiscale Materials Modeling [25], [61] and at the International Workshop on DD simulations [62].
This work is developed in the framework of the ANR OPTIDIS.

6.6.2. Co-design for scalable numerical algorithms in scientific applications

6.6.2.1. MHD instabilities edge localized modes

The last contribution of Xavier Lacoste's thesis deals with the integration of our work in JOREK, a production controlled plasma fusion simulation code from CEA Cadarache. We described a generic finite element oriented distributed matrix assembly and solver management API. The goal of this API is to optimize and simplify the construction of a distributed matrix which, given as an input to PaStiX, can improve the memory scaling of the application. Experiments exhibit that using this API we could reduce the memory consumption by moving to a distributed matrix input and improve the performance of the factorized matrix assembly by reducing the volume of communication. All this study is related to PaStiX integration inside JOREK but the same API could be used to produce a distributed assembly for another solver or/and another finite elements based simulation code.

6.6.2.2. Turbulence of plasma particules inside a tokamak

Concerning the GYSELA global non-linear electrostatic code, the efforts during the period have concentrated on predicting memory requirement and on the gyroaverage operator.

The Gysela program uses a mesh of 5 dimensions of the phase space (3 dimensions in configuration space and 2 dimensions in velocity space). On the large cases, the memory consumption already reaches the limit of the available memory on the supercomputers used in production (Tier-1 and Tier-0 typically). Furthermore, to implement the next features of Gysela (e.g. adding kinetic electrons in addition to ions), the needs of memory will dramatically increase, the main unknown will represents hundreds of TB. In this context, two tools were created to analyze and decrease the memory consumption. The first one is a tool that plots the memory consumption of the code during a run. This tool helps the developer to localize where the memory peak is located. The second tool is a prediction tool to compute the peak memory in offline mode (for production use mainly). A post processing stage combined with some specific traces generated on purpose during runtime allow the analysis of the memory consumption. Low-level primitives are called to generate these traces and to model memory consumption : they are included in the libMTM library (Modeling and Tracing Memory). Thanks to this work on memory consumption modeling, we have decreased the memory peak of the GYSELA code up to 50 % on a large case using 32,768 cores and memory scalability improvement has been shown using these tools up to 65k cores.

The main unknown of the Gysela is a distribution function that represents either the density of the guiding centers, either the density of the particles in a tokamak (depending of the location in the code). The switch between these two representations is done thanks to the gyroaverage operator. In the actual version of Gysela, the computation of this operator is achieved thanks to the so-called Padé approximation. In order to improve the precision of the gyroaveraging, a new implementation based on interpolation methods has been done (mainly by researchers from the Inria Tonus project-team and IPP Garching). We have performed the integration of this new implementation in GYSELA and also some parallel benchmarks. However, the new gyroaverage operator is approximatively 10 times slower than the original one. Investigations and optimizations on this operator are still a work in progress.
This work is carried on in the framework of Fabien Rozar's PhD in collaboration with CEA Cadarache.

6.6.2.3. SN Cartesian solver for nuclear core simulation

High-fidelity nuclear power plant core simulations require solving the Boltzmann transport equation. In discrete ordinate methods, the most computationally demanding operation of this equation is the sweep operation. Considering the evolution of computer architectures, we propose in this work, as a first step toward heterogeneous distributed architectures, a hybrid parallel implementation of the sweep operation on top of the generic task-based runtime system: PaRSEC. Such an implementation targets three nested levels of parallelism: message passing, multi-threading, and vectorization. A theoretical performance model was designed to validate the approach and help the tuning of the multiple parameters involved in such an approach. The proposed parallel implementation of the Sweep achieves a sustained performance of 6.1 Tflop/s, corresponding to 33.9% of the peak performance of the targeted supercomputer. This implementation compares favorably with state-of-art solvers such as PARTISN; and it can therefore serve as a building block for a massively parallel version of the neutron transport solver DOMINO developed at EDF.
Preliminary results have been presented at the international HPCC workshop on HPC-CFD in Energy/Transport Domains [50] in Paris. The main contribution will be presented at the international conference IPDPS 2015 [33] in Hyderabad.

6.6.2.4. 3D aerodynamics for unsteady problems with moving bodies

In the first part of our research work concerning the parallel aerodynamic code FLUSEPA, a first OpenMP-MPI version based on the previous one has been developped. By using an hybrid approach based on a domain decomposition, we achieved a faster version of the code and the temporal adaptive method used without bodies in relative motion has been tested successfully for real complex 3D-cases using up to 400 cores. Moreover, an asynchronous strategy for computing bodies in relative motion and mesh intersections has been developed and has been used for actual 3D-cases. A journal article (for JCP) to sum-up this part of the work is under redaction and a presentation at ISC at the "2nd International Workshop on High Performance Computing Simulation in Energy/Transport Domains" on July 2015 is scheduled.

This intermediate version exhibited synchronization problems for the aerodynamic solver due to the time integration used by the code. To tackle this issue, a task-based version over the runtime system StarPU is currently under development and evaluation. This year was mainly devoted to the realisation of this version. Task generation function have been designed in order to maximize asynchronism in execution. Those functions respect the data pattern access of the code and led to the refactorization of the actual kernels. A task-based version is now available for the aerodynamic solver and is available for both shared and distributed memory. This work will be presented as a poster during the SIAM CSE'15 conference and we are in the process to submit a paper in the Parallel CFD'15 conference.

The next steps will be to validate the correction of this task-based version and to work on the performance of this new version on actual cases. Later, the task description should be extended to the motion and intersection operations.
This work is carried on in the framework of Jean-Marie Couteyen's PhD in collaboration with Airbus Defence and Space Les Mureaux.

HIPERCOM2 Team

6. New Results

6.1. Highlights of the Year

- Hipercom 2 took part to the Inria-Industry meeting focusing on Telecommunications organized by Inria at Rocquencourt in November 2014. We presented a demonstration of the OCARI wireless sensor network.
- Hipercom 2 organized an Inria-DGA day "Software Defined Network (SDN) & MANET" at Paris in October 2014.

6.2. New Results about Wireless Sensor Networks

6.2.1. Node activity scheduling and routing in Wireless Sensor Networks

Participants: Cédric Adjih, Ichrak Amdouni, Pascale Minet.

The need to maximize network lifetime in wireless ad hoc networks and especially in wireless sensor networks requires the use of energy efficient algorithms and protocols. Motivated by the fact that a node consumes the least energy when its radio is in sleep state, we achieve energy efficiency by scheduling nodes activity. Nodes are assigned time slots during which they can transmit and they can turn off their radio when they are neither transmitting nor receiving. Compared to classical TDMA-based medium access scheme, spatial bandwidth use is optimized: non interfering nodes are able to share the same time slots, collisions are avoided and overhearing and interferences are reduced.

In 2014, we study the issue of delay optimization and energy efficiency in grid wireless sensor networks (WSNs). We focus on STDMA (Spatial Reuse TDMA)) scheduling, where a predefined cycle is repeated, and where each node has fixed transmission opportunities during specific slots (defined by colors). We assume a STDMA algorithm that takes advantage of the regularity of grid topology to also provide a spatially periodic coloring ("tiling" of the same color pattern). In this setting, the key challenges are: 1) minimizing the average routing delay by ordering the slots in the cycle 2) being energy efficient. Our work follows two directions: first, the baseline performance is evaluated when nothing specific is done and the colors are randomly ordered in the STDMA cycle. Then, we propose a solution, ORCHID that deliberately constructs an efficient STDMA schedule. It proceeds in two steps. In the first step, ORCHID starts form a colored grid and builds a hierarchical routing based on these colors. In the second step, ORCHID builds a color ordering, by considering jointly both routing and scheduling so as to ensure that any node will reach a sink in a single STDMA cycle. We study the performance of these solutions by means of simulations and modeling. Results show the excellent performance of ORCHID in terms of delays and energy compared to a shortest path routing that uses the delay as a heuristic. We also present the adaptation of ORCHID to general networks under the SINR interference model.

6.2.2. Time slot and channel assignment in multichannel Wireless Sensor Networks

Participants: Pascale Minet, Ridha Soua, Erwan Livolant.

Applying WSNs in industrial environment requires fast and reliable data gathering (or data convergecast). If packets are forwarded individually to the sink, it is called raw data convergecast. We resort to the multichannel paradigm to enhance the data gathering delay, the robustness against interferences and the throughput. Since some applications require deterministic and bounded convergecast delays, we target conflict free joint time slot and channel assignment solutions that minimize the schedule length. Such solutions allow nodes to save energy by sleeping in any slot where they are not involved in transmissions.

After a comprehensive survey on multichannel assignment protocols in wireless sensor networks, we study raw convergecast in multichannel wireless sensor networks (WSNs) where the sink may be equipped with multiple radio interfaces. We propose $Wave$,a simple, efficient and traffic-aware distributed joint channel and time slot assignment for raw convergecast. Our target is to minimize the data gathering delays and ensure that all packets transmitted in a cycle are delivered to the sink in this cycle, assuming no packet loss at the physical layer. We evaluate the number of slots needed to complete the convergecast by simulation and compare it to the optimal schedule and to a centralized solution. Simulations results indicate that our heuristic is not far from the optimal bound for raw convergecast. Unlike most previously published papers, $Wave$ does not suppose that all interfering links have been removed by channel allocation. In addition, $Wave$ is able to easily adapt to traffic changes. $Wave$ could be used to provide the schedule applied in the $802.15.4e$ TSCH based networks.

6.2.3. Optimized WSN Deployment

Participants: Ines Khoufi, Pascale Minet, Erwan Livolant.

This is a joint work with Telecom SudParis: Anis Laouiti.
We are witnessing the deployment of many wireless sensor networks in various application domains such as pollution detection in the environment, intruder detection at home, preventive maintenance in industrial process, monitoring of a temporary industrial worksite, damage assessment after a disaster.... Many of these applications require the full coverage of the area considered. With the full coverage of the area, any event occurring in this area is detected by at least one sensor node. In addition, the connectivity ensures that this event is reported to the sink in charge of analyzing the data gathered from the sensors and acting according to these data.

In the literature, many studies assume that this area is rectangular and adopt the classical deployment in triangular lattice that has been proved optimal. In real life, things are more complex. For instance, in an industrial worksite, the area to cover has an irregular shape with many edges and is not necessarily convex. Moreover, few papers take obstacles into account. Those that do assume that obstacles are constituted by a juxtaposition of rectangles that seems an unrealistic assumption. In real deployments, the shape of obstacles may be irregular. We distinguish two types of obstacles: the transparent ones like ponds in outdoor environment, or tables in an indoor site that only prevent the location of sensor nodes inside them; whereas the opaque obstacles like walls or trees prevent the sensing by causing the existence of hidden zones behind them: such zones may remain uncovered. Opaque obstacles are much more complex to handle than transparent ones and require the deployment of additional sensors to eliminate coverage holes. That is why we focus on the deployment of wireless sensor nodes in an arbitrary realistic area with an irregular shape, and with the presence of obstacles that may be opaque. Moreover, we propose a method that tends to minimize the number of sensor nodes needed to fully cover such an area.

Mobile robots can be used to deploy static wireless sensor nodes to achieve the coverage and connectivity requirements of the applications considered. Many solutions have been provided in the literature to compute the set of locations where the sensor nodes should be placed. We show how this set of locations can be used by a mobile robot to optimize its tour to deploy the sensor nodes to their right locations. In order to reduce both the energy consumed by the robot, its exposure time to a hostile environment, as well as the time at which the wireless network becomes operational, the optimal tour of the robot is this minimizing the delay. This delay must take into account not only the time needed by the robot to travel the tour distance but also the time spent in the rotations performed by the robot each time it changes its direction. This problem is called the Robot Deploying Sensor nodes problem, in short RDS. We first show how this problem differs from the well-known traveling salesman problem. We then propose an integer linear program formulation of the RDS problem. We propose various algorithms relevant to iterative improvement by exchanging tour edges, genetic approach and hybridization. The solutions provided by these algorithms are compared and their closeness to the optimal is evaluated in various configurations.

6.2.4. Sinks Deployment and Packet Scheduling for Wireless Sensor Networks

Participants: Nadjib Achir, Paul Muhlethaler.

The objective of this work is to propose an optimal deployment and distributed packet scheduling of multi-sink Wireless Sensors networks (WNSs). We start by computing the optimal deployment of sinks for a given maximum number of hops between nodes and sinks. We also propose an optimal distributed packet scheduling in order to estimate the minimum energy consumption. We consider the energy consumed due to reporting, forwarding and overhearing. In contrast to reporting and forwarding, the energy used in overhearing is difficult to estimate because it is dependent on the packet scheduling. In this case, we determine the lower-bound of overhearing, based on an optimal distributed packet scheduling formulation. We also propose another estimation of the lower-bound in order to simulate non interfering parallel transmissions which is more tractable in large networks. We note that overhearing largely predominates in energy consumption. A large part of the optimizations and computations carried out in this work are obtained using ILP formalization.

6.2.5. Security in wireless sensor networks

Participants: Selma Boumerdassi, Paul Muhlethaler.

Sensor networks are often used to collect data from the environment where they are located. These data can then be transmitted regularly to a special node called a *sink*, which can be fixed or mobile. For critical data (like military or medical data), it is important that sinks and simple sensors can mutually authenticate so as to avoid data to be collected and/or accessed by fake nodes. For some applications, the collection frequency can be very high. As a result, the authentication mechanism used between a node and a sink must be fast and efficient both in terms of calculation time and energy consumption. This is especially important for nodes which computing capabilities and battery lifetime are very low. Moreover, an extra effort has been done to develop alternative solutions to secure, authenticate, and ensure the confidentiality of sensors, and the distribution of keys in the sensor network. Specific researches have also been conducted for large-scale sensors. At present, we work on an exchange protocol between sensors and sinks based on low-cost shifts and xor operations.

6.2.6. Massive MIMO Cooperative Communications for Wireless Sensor Networks

Participants: Nadjib Achir, Paul Muhlethaler.

This work is a collaboration with Mérouane Debbah (Supelec, France).
The objective of this work is to propose a framework for massive MIMO cooperative communications for Wireless Sensor Networks. Our main objective is to analyze the performances of the deployment of a large number of sensors. This deployment should cope with a high demand for real time monitoring and should also take into account energy consumption. We have assumed a communication protocol with two phases: an initial training period followed by a second transmit period. The first period allows the sensors to estimate the channel state and the objective of the second period is to transmit the data sensed. We start analyzing the impact of the time devoted to each period. We study the throughput obtained with respect to the number of sensors when there is one sink. We also compute the optimal number of sinks with respect to the energy spent for different values of sensors. This work is a first step to establish a complete framework to study energy efficient Wireless Sensor Networks where the sensors collaborate to send information to a sink. Currently, we are exploring the multi-hop case.

6.2.7. Opportunistic routing cross-layer schemes for low duty-cycle wireless sensor networks

Participants: Mohamed Zayani, Paul Muhlethaler.

This is a joint work with Nadjib Aitsaadi from University of Paris 12.
The opportunistic aspect of routing is suitable with such networks where the topology is dynamic and protocols based on topological information become inefficient. Previous work initiated by Paul Muhlethaler and Nadjib Aitsaadi consisted in a geographical receiver-oriented scheme based on RI-MAC protocol (Receiver-Initiated MAC). This scheme is revised and a new contribution proposes to address the same problem with a sender-oriented approach. After scrutinising different protocols belonging to this classification, the B-MAC protocol is chosen to build a new opportunistic cross-layer scheme. Our choice is motivated by the ability of this protocol to provide to a sender the closest neighbor to the destination (typically a sink). In other words, such a scheme enables us to obtain shorter paths in terms of hops which would increase the efficiency of information delivery. In counterparts, as it relies on long preambles (property of B-MAC) to solicit all the neighborhood, it needs

larger delays and energy consumption (1% of active time). Nevertheless, this proposal remains interesting as the studied networks are dedicated to infrequent event detection and are not real time-oriented.

When we use BMAC with opportunistic routing, one main advantage is that there is no transmission when there is no event detected in the network in contrast to RI-MAC where beacons of awaking nodes are periodically sent. However, when an event occurs in the area monitored, the end-to-end delay to deliver the alert packet to the sink is much greater with BMAC than with RI-MAC. This may pose problem to some real-time applications. We have propose a scheme where, instead of sending a long preamble to gather all the neighbor nodes, the packet is directly sent. The acknowledgement of the packet allows tthe sender to know whether (or not) the progression towards the destination is sufficient. If it is not the case the packet is sent again. More neighbor node will be awaken and the progression towards the destination will be improved. The selection of the relay terminates when the progression towards the destination is above a given threshold. Actually this relaying scheme encompasses two levels of opportunism. The first level consists in selecting only the awake nodes, the second level consists in selecting the best nodes among the awake nodes. We can show that doing so only slightly increase the number of hops to reach the sink whereas the delay per hop is largely reduced. Thus the end-to-end is very significantly reduced and we still have the property that there is no transmission when there is no event detected in the network.

6.3. Cognitive Radio Networks

6.3.1. Multichannel time slot assignment in Cognitive Radio Sensor Networks

Participants: Ons Mabrouk, Pascale Minet, Ridha Soua, Ichrak Amdouni.

This is a joint work with Hanen Idoudi and Leila Saidane from ENSI, Tunisia.
The unlicensed spectrum bands become overcrowded causing an increased level of interference for current wireless sensor nodes. Cognitive Radio Sensor Networks (CRSNs) overcome this problem by allowing sensor nodes to access opportunistically the underutilized licensed spectrum bands. The sink assigns the spectrum holes to the secondary users (SUs). Therefore, it must rely on reliable information about the spectrum holes to protect the primary users (PUs). In 2013 we focused on the MultiChannel Time Slot Assignment problem (MC-TSA) in CRSN and proposed an Opportunistic centralized TIme slot assignment in COgnitive Radio sensor networks (OTICOR). This latter differs from the existing schemes in its ability to allow non-interfering cognitive sensors to access the same channel and time slot pair. OTICOR takes advantages of spatial reuse, multichannel communication and multiple radio interfaces of the sink. We proved through simulations that a smaller schedule length improves the throughput. Applying OTICOR, we show that, even in the presence of several PUs, the average throughput granted to SUs remains important. We also show how to get the best performances of OTICOR when the channel occupancy by PUs is known.

In 2014, we proposed two ways for the sink to determine the available channels and alert the SUs if an unexpected activity of PU occurs. Our objective is to design an algorithm able to detect the unexpected presence of PUs in the multi-hop network while maximizing the throughput. To achieve our goal, we propose an optimized version of our previous scheduling algorithm Opportunistic centralized TIme slot assignment in COgnitive Radio sensor networks (OTICOR). This algorithm takes advantage of the slots dedicated to the control period by allowing noninterfering cognitive sensors to access the control/data channel and time slot pair. We shown through simulations that using the control period for data transmission minimizes the schedule length and maximizes the throughput.

6.4. Mobile ad hoc and mesh networks

6.4.1. Development and implementation of a network coding module for NS3

Participants: Cédric Adjih, Ichrak Amdouni, Hana Baccouch.

DragonNet is a complete modular solution of network coding. This solution is responsible of coding, decoding, maintaining necessary information and the associated signaling. It is designed to be extensible. A variant of DragonNet was specified for wireless sensor networks and implemented.

As a follow-up to the ADT MOBSIM (and the previous module EyWifi), DragonNet was also integrated as a module for the NS-3 simulation tool.

6.4.2. Optimized Broadcast Scheme for Mobile Ad hoc Networks

Participants: Nadjib Achir, Paul Muhlethaler.

The main objective is to select the most appropriate relay nodes according to a given cost function. Basically, after receiving a broadcast packet each potential relay node computes a binary code according to a given cost function. Then, each node starts a sequence of transmit/listen intervals following this code. In other words, each 0 corresponds to a listening interval and each 1 to a transmit interval. During this active acknowledgment signaling period, each receiver applies the following rule: if it detects a signal during any of its listening intervals, it quits the selection process, since a better relay has also captured the packet. Finally, we split the transmission range into several sectors and we propose that all the nodes within the same sector use the same CDMA orthogonal spreading codes to transmit their signals. The CDMA codes used in two different sectors are orthogonal, which guarantees that the packet is broadcast in all possible directions. The obtained results demonstrate that our approach outperforms the classical flooding by increasing the delivery ratio and decreasing the number of required relays and thus the energy-cost.

6.5. Learning for an efficient and dynamic management of network resources and services

6.5.1. Learning in wireless sensor networks

Participants: Dana Marinca, Nesrine Ben Hassine, Pascale Minet, Selma Boumerdassi.

To guarantee an efficient and dynamic management of network resources and services we intend to use a powerful mathematical tool: prediction and learning from prediction. Prediction will be concerned with guessing the short-term, average-term and long-term evolution of network or network components state, based on knowledge about the past elements and/or other available information. Basically, the prediction problem could be formulated as follows: a forecaster observes the values of one or several metrics giving indications about the network state (generally speaking the network represents the environment). At each time t, before the environment reveals the new metric values, the forecaster predicts the new values based on previous observations. Contrary to classical methods where the environment evolution is characterized by stochastic process, we suppose that the environment evolution follows an unspecified mechanism, which could be deterministic, stochastic, or even adaptive to a given behavior. The prediction process should adapt to unpredictable network state changes due to its non-stationary nature. To properly address the adaptivity challenge, a special type of forecasters is used: the experts. These experts analyse the previous environment values, apply their own computation and make their own prediction. The experts predictions are given to the forecaster before the next environment values are revealed. The forecaster can then make its own prediction depending on the experts' "advice". The risk of a prediction may be defined as the value of a loss function measuring the discrepancy between the predicted value and the real environment value. The principal notion to optimize the behavior of the forecasters is the regret, seen as a difference between the forecaster's accumulated loss and that of each expert. To optimize the prediction process means to construct a forecasting strategy that guarantees a small loss with respect to defined experts. Adaptability of the forecaster is reflected in the manner in which it is able to follow the better expert according to the context.

In 2014, we applied on-line learning strategies to predict the quality of a wireless link in a WSN, based on the LQI metric and take advantage of wireless links with the best possible quality to improve the packet delivery rate. We model this problem as a forecaster prediction game based on the advice of several experts. The forecaster learns on-line how to adjust its prediction to better fit the environment metric values. A forecaster estimates the LQI value using the advice of experts. The model we propose learns on-line how to adapt to dynamic changes of the environment to compute efficient predictions. It presents a very good reactivity and adaptability. The simulations using traces collected in a real WSN based on the IEEE 802.15.4 standard have shown that the past time-windows which are effective for the prediction should have medium durations, about

200-400ms. The time windows durations less than 200ms do not give a good prediction, while durations larger than 400ms are efficient only in low variations environment. We note that these results strongly depend on the real traces, but the great advantage of the model is that it is self-adaptive to input traces profile. In this context, because of data normalization, the impact of loss functions is limited: entropy and square loss functions seem to give better and more stable predictions. Also, the experts prediction method should be adapted to traces profile. For low variation environment values, the average on past time windows is a good approximation. For high variation environment, a method predicting smoothed values close to minimum real values is more appropriate. Hence, the predicted values will be stabilized around the low values, avoiding estimations varying too much. Simulation results also show that for both types of experts (AMW and SES), the best expert depends on the phase considered. This is the reason why a forecaster is needed. Furthermore, the predictions of the EWA forecaster using SES experts are shown to be reactive and accurate. This combination minimizes the cumulated loss regarding the real LQI values, compared with any other combination such as EWA-AMW, BE-AMW and BE-SES, given by decreasing performance order.

6.5.2. Prediction and energy efficiency for datacenters

Participants: Dana Marinca, Nesrine Ben Hassine, Pascale Minet, Selma Boumerdassi.

The exponential development of Information and Communication Technologies (ICT) have led to an over consumption of services and data shared in networks. From computing in companies to unified communications through social networks and Internet of Things, the use of ICT a reach the highest level ever. The complexity involved by these different services reveals the limits of computing in companies and leads a majority of organisms to partially or completely host the management of there information system in data centers. The latter are larger and larger and are composed of buildings containing powerful computing equipments and air-conditionning systems. Data centers require a huge amount of energy. As an example, in 2014, the electric consumption of all date centers will be larger than 42 TWh, and after 2020 the CO2 production will be larger then 1.27 GTons, ie. more than the aeronautic industry (GeSI SMARTer 2020 report). These "frightening" figures led the research community to work on the management of energy consumption. Several tracks have been explored, among which the optimization of computation and load balancing of servers. At present, we work on tools dedicated to traffic prediction, thus allowing a better management of servers. Our work consists in modeling the traffic specific to data centers and apply different statistical prediction methods.

6.6. Vehicular Ad hoc NETworks (VANETs)

6.6.1. Congestion Control in VANETs

Participants: Paul Muhlethaler, Anis Laouiti.

We have reviewed the schemes of Congestion Control in VANETs for safety messages. The solutions proposed are: to adapt the generation rate, to adapt the transmission power or to adapt the carrier sense threshold. Some mechanisms employ different states depending on the channel load. Some other schemes use recursive adaptation of their parameters (e.g. LIMERIC). According to a few studies the recursive adaptation system provide a better adaptation of the VANET to the channel load. We will study how the transmission rate and the carrier sense threshold (or transmission power) can be best adapted in order to send CAM: Car Awareness Messages with the highest rate and to the furthest vehicles while maintaining the total load below a given threshold. We will also study the better combination of transmission rate and the carrier sense threshold for the CAM.

6.6.2. TDMA schemes for VANETs

Participants: Mohamed Hadded, Paul Muhlethaler, Anis Laouiti.

This is a joint work with Leila Saidane and Rachid Zabrouba from ENSI (Tunisia).

Vehicular Ad-hoc NETworks (VANETs) help improving traffic safety and efficiency. Each vehicle can exchange information to inform other vehicles about the current status of the traffic flow or a dangerous situation such as an accident. Road safety and traffic management applications require a reliable communication scheme with minimal transmission collisions, which thus increases the need for an efficient Medium Access Control (MAC) protocol. However, the MAC in a vehicular network is a challenging task due to the high speed of the nodes, frequent changes in topology, the lack of an infrastructure, and various QoS requirements. Recently several Time Division Multiple Access (TDMA)-based medium access control protocols have been proposed for vehicular ad hoc networks in an attempt to ensure that all the vehicles have enough time to send safety messages without collisions and reducing end-to-end delay and packet loss rate. We have identified the reasons for using the collision-free medium access control paradigm in VANETs. We have then presented a novel topology-based classification and we provide an overview of TDMA-based MAC protocols that have recently been proposed for VANETs. We have focus on the characteristics of these protocols as well as their benefits and limitations. Finally we have given a qualitative comparison, and we have discussed some open issues that need to be tackled in future studies to improve the performance of TDMA-based MAC protocols for vehicle to vehicle V2V communication.

HYBRID Project-Team

6. New Results

6.1. Highlights of the Year

- Paper [22] from Merwan Achibet, Maud Marchal, Ferran Argelaguet and Anatole Lécuyer received the "Best Paper Award" at IEEE Symposium on 3D User Interfaces 2014 (IEEE 3DUI'14).
- Paper [26] from Jean-Baptiste Barreau, Valérie Gouranton received the "Third Best Poster Award" at International Conference on Cultural Heritage 2014.

BEST PAPERS AWARDS :

[22] **IEEE Symposium on 3D User Interfaces**. M. ACHIBET, M. MARCHAL, F. ARGELAGUET SANZ, A. LÉCUYER.

6.2. 3D User Interfaces

6.2.1. 3D manipulation of virtual objects

Evaluation of Direct Manipulation using Finger Tracking for Complex Tasks in an Immersive Cube Maud Marchal, Collaboration with REVES

We have proposed a solution for interaction using finger tracking in a cubic immersive virtual reality system (or immersive cube) [13]. Rather than using a traditional wand device, users can manipulate objects with fingers of both hands in a close-to-natural manner for moderately complex, general purpose tasks. Our solution couples finger tracking with a real-time physics engine, combined with a heuristic approach for hand manipulation, which is robust to tracker noise and simulation instabilities. A first study has been performed to evaluate our interface, with tasks involving complex manipulations, such as balancing objects while walking in the cube. The users finger-tracked manipulation was compared to manipulation with a 6 degree-of-freedom wand (or flystick), as well as with carrying out the same task in the real world. Users were also asked to perform a free task, allowing us to observe their perceived level of presence in the scene. Our results show that our approach provides a feasible interface for immersive cube environments and is perceived by users as being closer to the real experience compared to the wand. However, the wand outperforms direct manipulation in terms of speed and precision. We conclude with a discussion of the results and implications for further research

A New Direct Manipulation Technique for Immersive 3D Virtual Environments Thi Thuong Huyen Nguyen, Thierry Duval, Collaboration with MIMETIC

We have introduced a new 7-Handle manipulation technique [38] for 3D objects in immersive virtual environments and its evaluation. The 7-Handle technique includes a set of seven points which are flexibly attached to an object. There are three different control modes for these points including configuration, manipulation and locking/ unlocking modes. We have conducted an experiment to compare the efficiency of this technique with the traditional 6-DOF direct manipulation technique in terms of time, discomfort metrics and subjective estimation for precise manipulations in an immersive virtual environment in two consecutive phases: an approach phase and a refinement phase. The statistical results showed that the completion time in the approach phase of the 7-Handle technique was significantly longer than the completion time of the 6-DOF technique. Nevertheless, we found a significant interaction effect between the two factors (the manipulation technique and the object size) on the completion time of the refinement phase. In addition, even though we did not find any significant differences between the two techniques in terms of intuitiveness, ease of use and global preference in the result of subjective data, we obtained a significantly better satisfaction feedback from the subjects for the efficiency and fatigue criteria.

A survey of plasticity in 3D user interfaces Jérémy Lacoche, Thierry Duval, Bruno Arnaldi, Collaboration with b<>com

Plasticity of 3D user interfaces [33] refers to their capabilities to automatically fit to a set of hardware and environmental constraints. This area of research has already been deeply explored in the domain of traditional 2D user interfaces. Besides, during the last decade, interest in 3D user interfaces has grown. Designers find with 3D user interfaces new ways to promote and interact with data, such as e-commerce websites, scientific data visualization, etc. Because of the wide variety of Virtual Reality (VR) and Augmented Reality (AR) applications in terms of hardware, data and target users, there is a real interest in solutions for automatic adaption in order to improve the user experience in any context while reducing the development costs. An adaptation is performed in reaction to different criteria defining a system such as the targeted hardware platform, the user's context and the structure and the semantic of the manipulated data. This adaptation can then impact the system in different ways, especially content presentation, interaction techniques modifications and eventually the current distribution of the system across a set of available devices. In [33] we present the state of the art about plastic 3D user interfaces. Moreover, we present well known methods in the field of 2D user interfaces that could become relevant for 3D user interfaces.

6.2.2. Navigating in virtual environments

Adaptive Navigation in Virtual Environments Ferran Argelaguet

Navigation speed for most navigation interfaces is still determined by rate-control devices (e.g. joystick). The interface designer is in charge of adjusting the range of optimal speeds according to the scale of the environment and the desired user experience. However, this approach is not valid for complex environments (e.g. multi-scale environments). Optimal speeds might vary for each section of the environment, leading to non-desired side effects such as collisions or simulator sickness. Thereby, we proposed a speed adaptation algorithm [24] based on the spatial relationship between the user and the environment and the user's perception of motion. The computed information is used to adjust the navigation speed in order to provide an optimal navigation speed and avoid collisions. Two main benefits of our approach is firstly, the ability to adapt the navigation speed in multi-scale environments and secondly, the capacity to provide a smooth navigation experience by decreasing the jerkiness of described trajectories. The evaluation showed that our approach provides comparable performance as existing navigation techniques but it significantly decreases the jerkiness of described trajectories

6.2.3. Novel pseudo-haptic based interfaces

Toward "Pseudo-Haptic Avatars": Modifying the Visual Animation ofSelf-Avatar Can Simulate the Perception of Weight Lifting, Ferran Argelaguet, Anatole Lécuyer, Collaboration with MIMETIC

We have studied how the visual animation of a self-avatar can be artificially modified in real-time in order to generate different haptic perceptions [18]. In our experimental setup, participants could watch their self-avatar in a virtual environment in mirror mode while performing a weight lifting task. Users could map their gestures on the self-animated avatar in real-time using a Kinect. We introduce three kinds of modification of the visual animation of the self-avatar according to the effort delivered by the virtual avatar: 1) changes on the spatial mapping between the user's gestures and the avatar, 2) different motion profiles of the animation ,and 3) changes in the posture of the avatar (upper-body inclination). The experimental task consisted of a weight lifting task in which participants had to order four virtual dumbbells according to their virtual weight. The user had to lift each virtual dumbbells by means of a tangible stick, the animation of the avatar was modulated according to the virtual weight of the dumbbell. The results showed that the altering the spatial mapping delivered the best performance. Nevertheless, participants globally appreciated all the different visual effects. Our results pave the way to the exploitation of such novel techniques in various VR applications such as sport training, exercise games, or industrial training scenarios in single or collaborative mode.

The Virtual Mitten: A Novel Interaction Paradigm for Visuo-Haptic Manipulation of Objects Using Grip Force Merwan Achibet, Maud Marchal, Ferran Argelaguet, Anatole Lécuyer

We have proposed a novel visuo-haptic interaction paradigm called the "Virtual Mitten" [22] for simulating the 3D manipulation of objects. Our approach introduces an elastic handheld device that provides a passive haptic feedback through the fingers and a mitten interaction metaphor that enables to grasp and manipulate objects. The grasping performed by the mitten is directly correlated with the grip force applied on the elastic device and a supplementary pseudo-haptic feedback modulates the visual feedback of the interaction in order to simulate different haptic perceptions. The Virtual Mitten allows natural interaction and grants users with an extended freedom of movement compared with rigid devices with limited workspaces. Our approach has been evaluated within two experiments focusing both on subjective appreciation and perception. Our results show that participants were able to well perceive different levels of effort during basic manipulation tasks thanks to our pseudo-haptic approach. They could also rapidly appreciate how to achieve different actions with the Virtual Mitten such as opening a drawer or pulling a lever. Taken together, our results suggest that our novel interaction paradigm could be used in a wide range of applications involving one or two-hand haptic manipulation such as virtual prototyping, virtual training or video game.

Figure 2. The Virtual Mitten : Visuo-haptic manipulation as enabled by our novel approach called the "Virtual Mitten". Each hand holds an elastic device to control a virtual mitten (in grey) enabling the grasp of virtual objects.

Collaborative Pseudo-Haptics: Two-User Stiffness Discrimination Based on Visual Feedback Ferran Argelaguet, Takuya Sato, Thierry Duval, Anatole Lécuyer, Collaboration with Tohoku University Research Institute of Electrical Communication

We have explored how the concept of pseudo-haptic feedback can be introduced in a collaborative scenario [25]. A remote collaborative scenario in which two users interact with a deformable object is presented. Each user, through touch-based input, is able to interact with a deformable virtual object displayed in a standard display screen. The visual deformation of the virtual object is driven by a pseudo-haptic approach taking into account both the user in-put and the simulated physical properties. Particularly, we investigated stiffness perception. In order to validate our approach, we tested our system in a single and two-user configuration. The results showed that users were able to discriminate the stiffness of the virtual object in both conditions with a comparable performance. Thus, pseudo-haptic feedback seems a promising tool for providing multiple users with physical information related to other users' interactions.

6.2.4. Sound and virtual reality

Sonic interaction with a virtual orchestra of factory machinery Florian Nouviale, Valérie Gouranton, collaboration with Ronan Gaugne (IMMERSIA) and LIMSI

We have conceived an immersive application where users receive sound and visual feedbacks on their interactions with a virtual environment. In this application, the users play the part of conductors of an orchestra of factory machines since each of their actions on interaction devices triggers a pair of visual and audio responses. Audio stimuli were spatialized around the listener. The application was exhibited during the 2013 Science and Music day and designed to be used in a large immersive system with head tracking, shutter glasses and a 10.2 loudspeaker configuration [43].

Audio-Visual Attractors for Capturing Attention to the Screens When Walking in CAVE Systems Ferran Argelaguet, Valérie Gouranton, Anatole Lécuyer, collaboration with Aalborg University

In four-sided CAVE-like VR systems, the absence of the rear wall has been shown to decrease the level of immersion and can introduce breaks in presence. We have therefore investigated to which extent user's attention can be driven by visual and auditory stimuli in a four-sided CAVE-like system [32]. An experiment was conducted in order to analyze how user attention is diverted while physically walking in a virtual environment, when audio and/or visual attractors are present. The four sided CAVE used in the experiment allowed to walk up to 9m in straight line. An additional key feature in the experiment is the fact that auditory feedback was delivered through binaural audio rendering techniques via non-personalized head related transfer functions (HRTFs). The audio rendering was dependent on the user's head position and orientation, enabling localized sound rendering. The experiment analyzed how different "attractors" (audio and/or visual, static or dynamic) modify the user's attention. The results of the conducted experiment show that audio-visual attractors are the most efficient attractors in order to keep the user's attention toward the inside of the CAVE. The knowledge gathered in the experiment can provide guidelines to the design of virtual attractors in order to keep the attention of the user and avoid the "missing wall".

6.2.5. Experiencing the past in virtual reality

Immersia, an open immersive infrastructure: doing archaeology in virtual reality Valérie Gouranton, Bruno Arnaldi, collaboration with MIMETIC and Ronan Gaugne (IMMERSIA)

We have first studied the mutual enrichment between archaeology and virtual reality [16]. To do so, we are considering Immersia, our open high-end platform dedicated to research on immersive virtual reality and its usages. Immersia is a node of the european project Visionair that offers an infrastructure for high level visualisation facilities open to research communities across Europe. In Immersia, two projects are currently active on the theme of archaeology. One is relative to the study of the Cairn of Carn, with the Creaah, a pluridisciplinary research laboratory of archeology and archeosciences, and one on the reconstitution of the gallo-roman villa of Bais, with the French institute INRAP.

Virtual reality tools for the West Digital Conservatory of Archaeological Heritage Jean-Baptiste Barreau, Valérie Gouranton, collaboration with Ronan Gaugne (IMMERSIA) and INRAP

In continuation of the 3D data production work made by the WDCAH (West Digital Conservatory of Archaeological Heritage), the use of virtual reality tools allows archaeologists to carry out analysis and understanding research about their sites. We have then focused on the virtual reality services proposed to archaeologists in the WDCAH, through the example of two archaeological sites, the Temple de Mars in Corseul and the Cairn of Carn Island [27].

Preservative Approach to Study Encased Archaeological Artefacts Valérie Gouranton, Bruno Arnaldi, collaboration with Ronan Gaugne (IMMERSIA) and INRAP

We have proposed a workflow based on a combination of computed tomography, 3D images and 3D printing to analyse different archaeological material dating from the Iron Age, a weight axis, a helical piece, and a fibula [39]. This workflow enables a preservative analysis of the artefacts that are unreachable because encased either in stone, corrosion or ashes. Computed tomography images together with 3D printing provide a rich toolbox for archaeologist work allowing to access a tangible representation of hidden artefacts. These technologies are combined in an efficient, affordable and accurate workflow compatible with preventive archaeology constraints.

Combination of 3D Scanning, Modeling and Analyzing Methods around the Castle of Coatfrec Reconstitution Jean Baptiste Barreau, Valérie Gouranton, collaboration with Ronan Gaugne (IMMERSIA) and INRAP

The castle of Coatfrec is a medieval castle in Brittany constituting merely a few remaining ruins currently in the process of restoration. Beyond its great archeological interest, it has become, over the course of the last few years, the subject of experimentation in digital archeology. Methods of 3D scanning were used in order to gauge comparisons between the remaining structures and their absent hypothetical ones, resulting in the first

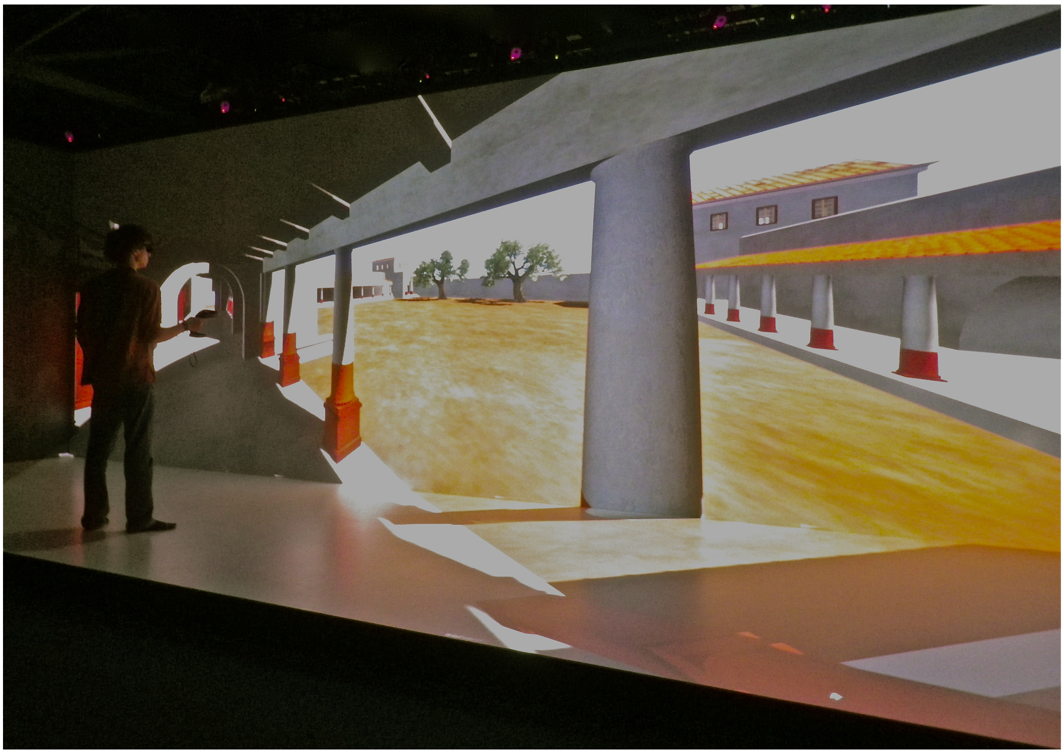


Figure 3. Virtual visit of the "Temple of Mars" in the Immersia room.

quantitative results of its kind. We have applied these methods and presented the subsequent results obtained using these new digital tools [26].

Ceramics Fragments Digitization by Photogrammetry, Reconstructions and Applications Jean Baptiste Barreau, Valérie Gouranton, collaboration with Ronan Gaugne (IMMERSIA) and INRAP

We have studied an application of photogrammetry on ceramic fragments from two excavation sites located north-west of France [28]. The restitution by photogrammetry of these different fragments allowed reconstructions of the potteries in their original state or at least to get to as close as possible. We used the 3D reconstructions to compute some metrics and to generate a presentation support by using a 3D printer. This work is based on affordable tools and illustrates how 3D technologies can be quite easily integrated in archaeology process with limited financial resources.

6.3. Physical simulation and multisensory feedback

6.3.1. Physically-based simulation and collision detection

Fast collision detection for fracturing rigid bodies Loeiz Glondu, Maud Marchal

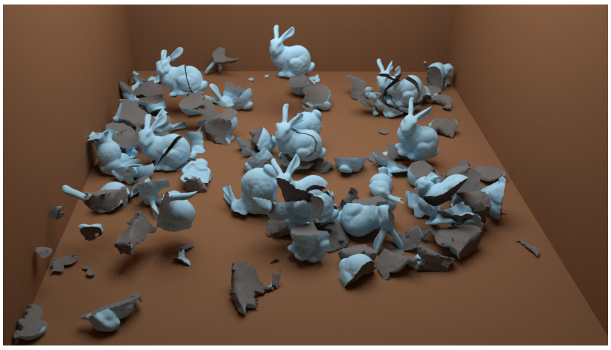


Figure 4. Fast collision detection during real-time brittle fracture simulation.

In complex scenes with many objects, collision detection plays a key role in the simulation performance. This is particularly true in fracture simulation for two main reasons. One is that fracture fragments tend to exhibit very intensive contact, and the other is that collision detection data structures for new fragments need to be computed on the fly. In [17], we present novel collision detection algorithms and data structures for real-time simulation of fracturing rigid bodies. We build on a combination of well-known efficient data structures, namely distance fields and sphere trees, making our algorithm easy to integrate on existing simulation engines. We propose novel methods to construct these data structures, such that they can be efficiently updated upon fracture events and integrated in a simple yet effective self-adapting contact selection algorithm. Altogether, we drastically reduce the cost of both collision detection and collision response. We have evaluated our global solution for collision detection on challenging scenarios, achieving high frame rates suited for hard real-time applications such as video games or haptics. Our solution opens promising perspectives for complex fracture simulations involving many dynamically created rigid objects.

This work was achieved in collaboration with Miguel Otaduy and Sara Schvartzman (URJC Madrid, Spain) and Georges Dumont (MIMETIC team).

Collision detection: broad phase adaptation from multi-core to multi-GPU architecture Bruno Arnaldi, Valérie Gouranton

We have presented several contributions on the collision detection optimization centered on hardware performance. We focus on the first step (Broad-phase) and propose three new ways of parallelization of the well-known Sweep and Prune algorithm. We first developed a multi-core model takes into account the number of available cores. Multi-core architecture enables us to distribute geometric computations with use of multi-threading. Critical writing section and threads idling have been minimized by introducing new data structures for each thread. Programming with directives, like OpenMP, appears to be a good compromise for code portability. We then proposed a new GPU-based algorithm also based on the "Sweep and Prune" that has been adapted to multi-GPU architectures. Our technique is based on a spatial subdivision method used to distribute computations among GPUs. Results show that significant speed-up can be obtained by passing from 1 to 4 GPUs in a large-scale environment [12].

Real-time tracking of deformable target in ultrasound images Maud Marchal

In several medical applications such as liver or kidney biopsies, an anatomical region needs to be continuously tracked during the intervention. When using ultrasound (US) image modality, tracking soft tissues remains challenging due to the deformations caused by physiological motions or medical instruments, combined with the generally weak quality of the images. In order to overcome the previous limitation, different techniques based on physical model have been proposed in the literature. In [41], we propose an approach for tracking deformable target within 2D US images based on a physical model driven by smooth displacement field obtained from dense information. This allows to take into account highly localized deformation in the US images.

This work was achieved in collaboration with Lucas Royer and Alexandre Krupa (Lagadic team), Anthony Le Bras (CHU Rennes) and Guillaume Dardenne (IRT B-Com).

6.3.2. Multimodal feedback

Stereoscopic Rendering of Virtual Environments with Wide Field-of-Views up to 360 Jérôme Ardouin, Anatole Lécuyer, Maud Marchal

We propose a novel approach [23] for stereoscopic rendering of virtual environments with a wide Field-of-View (FoV) up to 360. Handling such a wide FoV implies the use of non-planar projections and generates specific problems such as for rasterization and clipping of primitives. We propose a novel pre-clip stage specifically adapted to geometric approaches for which problems occur with polygons spanning across the projection discontinuities. Our approach integrates seamlessly with immersive virtual reality systems as it is compatible with stereoscopy, head-tracking, and multi-surface projections. The benchmarking of our approach with different hardware setups shows that it is well compliant with real-time constraints, and capable of displaying a wide range of FoVs. Thus, our geometric approach could be used in various VR applications in which the user needs to extend the FoV and apprehend more visual information.

This work was achieved in collaboration with Eric Marchand (Lagadic team).

A survey on bimanual haptics Anatole Lécuyer, Maud Marchal, Anthony Talvas.

When interacting with virtual objects through haptic devices, most of the time only one hand is involved. However, the increase of computational power, along with the decrease of device costs, allow more and more the use of dual haptic devices. The field which encompasses all studies of the haptic interaction with either remote or virtual environments using both hands of the same person is referred to as bimanual haptics. It differs from the common unimanual haptic field notably due to specificities of the human bimanual haptic system, e.g. the dominance of the hands, their differences in perception and their interactions at a cognitive level. These specificities call for adapted solutions in terms of hardware and software when applying the use of two hands to computer haptics. In [21], we review the state of the art on bimanual haptics, encompassing the human factors in bimanual haptic interaction, the currently available bimanual haptic devices, the software solutions for two-handed haptic interaction, and the existing interaction techniques.

Haptic cinematography Fabien Danieau, Anatole Lécuyer

Haptics, the technology which brings tactile or force-feedback to users, has a great potential for enhancing movies and could lead to new immersive experiences. In [14] we introduce *Haptic Cinematography* which presents haptics as a new component of the filmmaker's toolkit. We propose a taxonomy of haptic effects and we introduce novel effects coupled with classical cinematographic motions to enhance video viewing experience. More precisely we propose two models to render haptic effects based on camera motions: the first model makes the audience feel the motion of the camera and the second provides haptic metaphors related to the semantics of the camera effect. Results from a user study suggest that these new effects improve the quality of experience. Filmmakers may use this new way of creating haptic effects to propose new immersive audiovisual experiences.

This work was achieved in collaboration with Marc Christie (MIMETIC team), Julien Fleureau, Philippe Guillotel and Nicolas Mollet (Technicolor).

6.4. Collaborative Virtual Environments

6.4.1. Collaborative virtual environments for training

Collaborative Virtual Training with Physical and Communicative Autonomous Agents Thomas Lopez, Valérie Gouranton, Florian Nouviale, Rozenn Bouville-Berthelot, Bruno Arnaldi

Virtual agents are a real asset in collaborative virtual environment for training (CVET) as they can replace missing team members. Collaboration between such agents and users, however, is generally limited. We presented a whole integrated model of CVET focusing on the abstraction of the real or virtual nature of the actor to define a homogenous collaboration model. First, we defined a new collaborative model of interaction. This model notably allows to abstract the real or virtual nature of a teammate. Moreover, we proposed a new role exchange approach so that actors can swap their roles during training. The model also permits the use of physically based objects and characters animation to increase the realism of the world. Second, we design a new communicative agent model, which aims at improving collaboration with other actors using dialog to coordinate their actions and to share their knowledge. Finally, we evaluated the proposed model to estimate the resulting benefits for the users and we show that this is integrated in existing CVET applications [20].

Figure 5. Collaborative virtual environment for training (CVET) as experienced in our Immersia VR room

Exchange of avatars : Toward a better perception and understanding Thomas Lopez, Rozenn Bouville-Berthelot, Florian Nouviale, Valérie Gouranton, Bruno Arnaldi

The exchange of avatars, i.e. the actual fact of changing once avatar with another one, is a promising trend in multi-actor virtual environments. It provides new opportunities for users, such as controlling a different avatar for a specific action, retrieving knowledge belonging to a particular avatar, solving conflicts and deadlocks situations or even helping another user. Virtual Environments for Training are especially affected by this trend as a specific role derived from a scenario is usually assigned to a unique avatar. Despite the increasing use of avatar exchange, users' perception and understanding of this mechanism have not been studied. We propose two complementary user-centered evaluations that aim at comparing several representations for the exchange of avatars; these are termed exchange metaphors. Our first experiment focuses on the perception of an exchange by a user who is not involved in the exchange, and the second experiment analyzes the perception of an exchange triggered by the user. Results show that the use of visual feedback globally aids better understanding of the exchange mechanism in both cases. Our first experiment suggests, however, that visual feedback is less efficient than a simple popup notification in terms of task duration. In addition, the second experiment shows that much simpler metaphors with no visual effect are generally preferred because of their efficiency [19].

Figure 6. Example of metaphor used in "Exchange of Avatars": the blue expert exchanges avatar here with the red expert.

An interaction abstraction model for seamless avatar exchange in CVET Rozenn Bouville-Berthelot, Thomas Lopez, Florian Nouviale, Valérie Gouranton, Bruno Arnaldi

Collaboration and interaction between users and virtual humans in virtual environments is a crucial challenge, notably for Collaborative Virtual Environments for Training (CVET). A training procedure, indeed, often involves several actors: trainees, teammates and many times a trainer. Yet, a major benefit of CVET is to propose to users to be trained even if the required number of person needed by the procedure is not available. Therefore, almost every CVET use autonomous virtual humans to replace the missing person. We have proposed to improve the effective collaboration between users and virtual humans involved in a complex task within CVET. Using an entity called the "Shell", we are able to wrap the features common to both users and virtual humans. It gives us an abstraction level to pool the management of the main processes useful to control an avatar, interact with the environment and gather knowledge from a CVET. Besides, the Shell allows seamless exchange of avatars during a procedure. Thanks to the Shell, the exchange can be carried out at any time during a task while preserving all the data associated to a role in a procedure [29].

#SEVEN: a Sensor Effector Based Scenarios Model for Driving Collaborative Virtual Environment Guillaume Claude, Valérie Gouranton, Rozenn Bouville-Berthelot, Bruno Arnaldi

We introduced #SEVEN, a sensor effector model that enables the execution of complex scenarios for driving Virtual Reality applications. #SEVEN is based on an enhanced Petri net model which is able to describe and solve intricate event sequences. Our model also proposes several useful features for the design of collaborative scenarios for Collaborative Virtual Environments such as versatile roles and Activity Continuum. We also illustrate its usage it by describing a demonstrator that presents an implementation of our model [30].

Collaborative virtual environments for ergonomics: embedding the design engineer role in the loop Thierry Duval, collaboration with Charles Pontonnier and Georges Dumont (MIMETIC).

We have proposed to define the role and duties of a design engineer involved in a collaborative ergonomic design session supported by a 3D collaborative virtual environment. For example, such a session can be used to adapt the manual task an operator must achieve in the context of an industrial assembly line. We first presented the interest of such collaborative sessions. Then we presented a related work explaining the need of proper 3DCVE and metaphors to obtain efficient collaborative ergonomic design sessions. Then we proposed a use case highlighting the type of metaphor such engineers need to have to be efficient in such a framework [40].

6.4.2. Collaborative virtual environments and awareness

Improving Awareness for 3D Virtual Collaboration by Embedding the Features of Users' Physical Environments and by Augmenting Interaction Tools with Cognitive Feedback Cues Thierry Duval, Thi Thuong Huyen Nguyen, Valérie Gouranton, collaboration with MimeTic

The feeling of presence is essential for efficient interaction within Virtual Environments (VEs). When a user is fully immersed within a VE through a large immersive display system, his/her feeling of presence can be altered because of disturbing interactions with his/her physical environment, such as collision with hardware parts of the system or loss of tracking. This alteration can be avoided by taking into account the physical features of the user as well as those of the system hardware and embedding them in the VE. Moreover, the 3D abstract representation of these physical features can also be useful for collaboration between distant users because they can make a user aware of the physical limitations of the others he/she is collaborating with. We used the Immersive Interactive Virtual Cabin (IIVC) model to obtain this virtual representation of the user's physical environment and we illustrated how this representation can be used in a collaborative navigation task in a VE. We also presented how we can add 3D representations of 2D interaction tools in order to cope with asymmetrical collaborative configurations, providing 3D cues for a user to understand the actions of the others even if he/she is not fully immersed in the shared VE [15].

From 3D Bimanual Toward Distant Collaborative Interaction Techniques: An Awareness Issue Morgan Le Chénechal, Thierry Duval, Valérie Gouranton, Bruno Arnaldi, collaboration with b<>com

CVE involve the use of complex interaction techniques based on specific collaborative metaphors. The design of these metaphors may be a difficult task because it has to deal with collaborative issues that came from sparse research areas (Human-Computer Interfaces, Human-Human Interactions, Networking, Physiology and Social Psychology). Metaphors for bimanual interactions have been developed for a while essentially because it is a widely spread area of interest for common tasks. Bimanual interactions involve the simultaneous use of both hands of the user in order to achieve a goal with better performances compared to uni-manual interactions thanks to a natural skill that is proprioception. This collaborative aspect could certainly be a helpful entry point in the design of efficient collaborative interaction techniques extended from improved bimanual metaphors. However, the proprioceptive sense cannot be considered in the same way, and additional features must be proposed to be able to collaborate efficiently. Thus, awareness is a key to let CVE be usable and the availability of collaborative feedbacks is essential to extend bimanual interactions toward collaborative ones. In this paper, we based our study on existing work on bimanual and collaborative interaction techniques trying to draw similarities between them. We emphasize common points between both fields that could be useful to better design both metaphors and awareness in CVE [34].

A survey of communication and awareness in collaborative virtual environments Thi Thuong Huyen Nguyen, Thierry Duval

In the domain of Collaborative Virtual Environments (CVEs), many virtual worlds, frameworks and techniques are built based on a specific and direct purpose. There is not a general and still good and efficient enough solution for all the collaborative systems. Depending on the purpose of the collaborative work, the techniques of interaction and of manipulation change from one application to another. Despite this difference between interaction techniques, they always benefit greatly from awareness features that help in explicating implicit knowledge related to one's own and others' working activities as well as to virtual workspace. In addition, people in CVEs also use communication channels to negotiate shared understandings of task goals, of task decomposition and of task progress. Therefore, awareness and communication are usually considered as "instruments" to complete collaborative tasks in the environment. However, few research work have been devoted to improving the awareness and the communication channels in CVEs for a better collaboration between users. We have studied the importance of awareness and communication in collaborative virtual environments. We have investigated different kinds of awareness which need to be carefully designed. We have discussed different communication means and how to cope with this diversity, so we can benefit from the availability of different peripheral devices and can find an effective communication means to work together. Finally, we have made some propositions to overcome these actual limitations of CVEs [37].

6.4.3. Collaborative virtual environments and software engineering

When model driven engineering meets virtual reality: feedback from application to the Collaviz framework Thierry Duval, collaboration with Arnaud Blouin and Jean-Marc Jézéquel (DIVERSE).

Despite the increasing use of 3D Collaborative Virtual Environments (3D CVE), their development is still a cumbersome task. The various concerns to consider (distributed system, 3D graphics, etc.) complexify the development as well as the evolution of CVEs. Software engineering recently proposed methods and tools to ease the development process of complex software systems. Among them, Model-Driven Engineering (MDE) considers models as first-class entities. A model is an abstraction of a specific aspect of the system under study for a specific purpose. MDE thus breaks down a complex system into as many models for different purposes, such as: generating code from models; building domain specific programming/modeling languages (DSL); generating tools such as graphical or textual editors. We have leveraged MDE for developing 3D CVEs. We showed how the Collaviz framework took benefits from a DSL we built. The benefits are multiple: 3D CVE designers can focus on the behavior of their virtual objects without bothering with distributed and graphics features; configuring the content of 3D CVEs and their deployment on various software and hardware platforms can be automated through code generation. We detailed the development process we propose and the experiments we conducted on Collaviz [31].

6.5. Brain-Computer Interfaces

6.5.1. Novel usages of BCI

Mind-Mirror: combining BCI and augmented reality to "see your brain in action in your own head", Anatole Lécuyer, Jonathan Mercier, Maud Marchal

Imagine you are facing a mirror, seeing at the same time both your real body and a virtual display of your brain in activity and perfectly superimposed to your real image "inside your real skull". We have introduced a novel augmented reality paradigm called "Mind-Mirror" which enables the experience of seeing "through your own head", visualizing your brain "in action and in situ" [36]. Our approach relies on the use of a semi-transparent mirror positioned in front of a computer screen. A virtual brain is displayed on screen

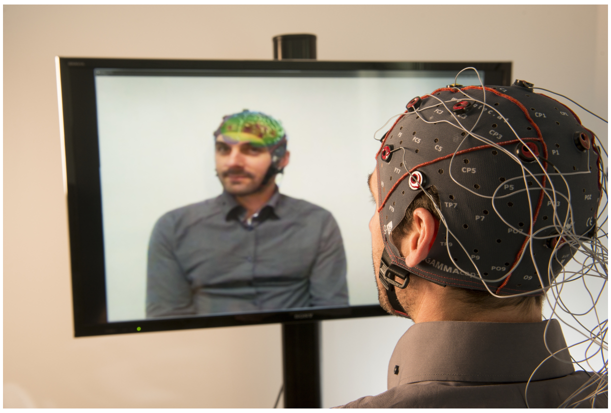


Figure 7. The Mind Mirror system : Visualization of brain activity in real-time using a mirror-based augmented reality setup and an EEG headset.

and automatically follows the head movements using an optical face-tracking system. The brain activity is extracted and processed in real-time with the help of an electroencephalography cap (EEG) worn by the user. A rear view is also proposed thanks to an additional webcam recording the rear of the user's head. The use of EEG classification techniques enables to test a Neurofeedback scenario in which the user can train and progressively learn how to control different mental states, such as "concentrated" versus "relaxed". The results of a user study comparing a standard visualization used in Neurofeedback to our approach showed that the Mind-Mirror could be successfully used and that the participants have particularly appreciated its innovation and originality. We believe that, in addition to applications in Neurofeedback and Brain-Computer Interfaces, the Mind-Mirror could also be used as a novel visualization tool for education, training or entertainment applications.

This work was achieved in collaboration with Fabien Lotte from POTIOC team (Inria-Bordeaux).

Using SSVEP-based BCI with 3D stereoscopic display Anatole Lécuyer

We have investigated the feasibility of dual-frequency Steady-State Visual Evoked Potential (SSVEP) stimulation using a 3-D display and stereoscopic glasses [44]. Dual-frequency stimulation allows for more targets to be created using a small number of frequencies, and stereoscopic vision offers a suitable medium for dual-frequency stimulation as the two views can be controlled independently. Participants were exposed to a repetitive visual stimulus flashing at different frequencies in the left and right views and the electroencephalography (EEG) trace was examined. Our results suggest that the two stimulation frequencies can still be evident in the SSVEP response. In addition, the participant ratings showed no significant differences in fatigue, annoyance, comfort or strangeness of the stimulation compared to conventional forms of stimulation. These results pave the way for further studies using stereoscopic dual-frequency stimulation and its potential for use in virtual reality and 3D videogames

This work was achieved in collaboration with Robert Leeb (EPFL, Switzerland).

Passive BCI and music Anatole Lécuyer

Passive brain–computer interfaces (passive BCI), also named implicit BCI, provide information from user mental activity to a computerized application without the need for the user to control his brain activity. We have proposed an overview of current research on passive BCIs in [45]. We have notably studied how they can be applied to the context of music creation, where they can provide novel information to adapt the music creation process, e.g., exploiting user mental concentration to adapt the music tempo.

6.5.2. BCI methodology

Which factors drive successful BCI skill learning? Anatole Lécuyer, Lorraine Perronnet

Brain-Computer Interfaces although very promising, suffer from a poor reliability. Rather than improving brain signal-processing alone, an interesting research direction is to guide users to learn BCI control mastery. Thus, we have inroduced a set of motivational and cognitive factors which could influence the learning process, and which should be considered to improve the global performance of BCI users [47]. We base our approach on Keller's integrative theory of motivation, volition, and performance, which combines motivational (affective) and cognitive factors, to explain what makes human users learn and perform efficiently, irrespectively of the task. These factors can guide the creation of learning environments, such as BCI training protocols.

This work was achieved in collaboration with Fabien Lotte and Christian Muhl (POTIOC team, Inria-Bordeaux), Moritz Grosse-Wentrup (MPI, Tuebingen), and Reinhold Scherer (TU Graz, Austria).

A methodological framework for applications combining BCI and VE Anatole Lécuyer

We have proposed a user-centred methodological framework [46] to guide design and evaluation of applications combining Brain-Computer Interface (BCI) and Virtual Environment (VE). Our framework is based on the contributions of ergonomics to ensure these applications are well suited for end-users. It provides methods, criteria and metrics to perform the phases of the human-centred design process aiming to understand the context of use, specify the user needs and evaluate the solutions in order to define design choices. Several ergonomic methods (e.g., interviews, longitudinal studies, user based testing), objective metrics (e.g., task success, number of errors) and subjective metrics (e.g., mark assigned to an item) are suggested to define and measure the usefulness, usability, acceptability, hedonic qualities, appealingness, emotions related to user experience, immersion and presence to be respected. The benefits and contributions of our user centred framework for the ergonomic design of applications combining BCI and VE were also discussed.

This work was achieved in collaboration with Fabien Lotte from POTIOC team (Inria-Bordeaux).

HYCOMES Team

6. New Results

6.1. Highlights of the Year

The main advances in 2014 of the Hycomes team have been as follows:

Causality analysis of hybrid systems with ordinary differential equations (ODE) We have proposed a causality analysis, in the form of a simple type system, rejecting hybrid programs with algebraic circuits — see section 6.2 .

An index theory of DAE hybrid systems with differential algebraic equations (DAE) We have proposed a conservative extension of the notion of differentiation index to hybrid systems with differential algebraic equations — see section 6.3 .

6.2. A Type-Based Analysis of Causality Loops In Hybrid Systems Modelers

Explicit hybrid systems modelers like Simulink / Stateflow allow for programming both discrete- and continuous-time behaviors with complex interactions between them. A key issue in their compilation is the static detection of algebraic or causality loops. Such loops can cause simulations to deadlock and prevent the generation of statically scheduled code. In [6] (also published as a deliverable of the Sys2Soft collaborative project [14], see 7.2), we address this issue for a hybrid modeling language that combines synchronous Lustre-like data-flow equations with Ordinary Differential Equations (ODEs). We introduce the operator last(x) for the left-limit of a signal x. This operator is used to break causality loops and permits a uniform treatment of discrete and continuous state variables. The semantics relies on non-standard analysis, defining an execution as a sequence of infinitesimally small steps. A signal is deemed causally correct when it can be computed sequentially and only progresses by infinitesimal steps outside of discrete events. The causality analysis takes the form of a simple type system. In well-typed programs, signals are proved continuous during integration and can be translated into sequential code for integration with off-the-shelf ODE solvers. The effectiveness of this system is illustrated with several examples written in Zélus, a Lustre-like synchronous language extended with hierarchical automata and ODEs.

6.3. On the index of multi-mode DAE Systems

Hybrid systems modelers exhibit a number of difficulties related to the mix of continuous and discrete dynamics and sensitivity to the discretization scheme. Modular modeling, where subsystems models can be simply assembled with no rework, calls for using Differential Algebraic Equations (DAE). In turn, DAE are strictly more difficult than ODE. In most modeling and simulation tools, before simulation can occur, sophisticated pre-processing is applied to DAE systems based on the notion of differentiation index. Graph based algorithms such as the one originally proposed by Pantelides [47] are efficient at finding the differentiation index of a DAE system, structurally (i.e., outside some exceptional values for the system parameters), solving the consistent initialisation problem and, transforming a DAE system into a statically scheduled system of ordinary differential equations (ODE) and implicit functions. The differentiation index for DAE explicitly relies on everything being differentiable. Therefore, extensions to hybrid systems must be done with caution — to our knowledge, no such extension exists, supported by a rigourous mathematical theory. In [8], we use non-standard analysis for this. Non-standard analysis formalizes differential equations as discrete step transition systems with an infinitesimal time basis. This allows to map hybrid DAE systems to difference Algebraic Equations (dAE), for which the notion of difference index can be used. The difference index of a dAE is an easy transposition of the differentiation index of a DAE, where forward shift in time (using a next() operator) replaces differentiation. We prove that the differentiation index of a DAE is structurally equal to the difference index of the dAE resulting from its non-standard interpretation. We can thus propose the difference index of the non-standard semantics of a hybrid DAE system, as a consistent extension of both the differentiation index of DAE and the difference index of dAE. It turns out that the index theory for (discrete time) dAE systems is interesting in itself and raises new issues. We have investigated graph based method similar to the Pantelides [47] algorithm for computing the difference index of a dAE.

6.4. A Unifying View of Loosely Time-Triggered Architectures

Cyber-Physical Systems require distributed architectures to support safety critical real-time control. Hermann Kopetz' Time-Triggered Architecture (TTA) has been proposed as both an architecture and a comprehensive paradigm for systems architecture, for such systems. TTA offers the programmer a logical discrete time compliant with synchronous programming, together with timing bounds. A clock synchronization protocol is required, unless the local clocks used themselves provide the recquired accuracy. To relax the strict requirements on synchronization imposed by TTA, Loosely Time-Triggered Architectures (LTTA) have been proposed. In LTTA, computation and communication units are all triggered by autonomous, unsynchronized, clocks. Communication media act as shared memories between writers and readers and communication is non blocking. This is at the price of communication artifacts (such as duplication or loss of data), which must be compensated for by using some "LTTA protocol". In [7] we have pursued our previous work by providing a unified presentation of the two variants of LTTA (token- and time-based), with simplified analyses. We compared these two variants regarding performance and robustness and we provide ways to combine them.

I4S Project-Team

6. New Results

6.1. Highlights of the Year

The team organized the 7th European Workshop on SHM in Nantes in July 2014 (http://ewshm2014.com) .

6.2. Analysis and control of systems

6.2.1. Optimal vibration damping of large structures

Participant: Dominique Siegert.

This paper deals with the theoretical and experimental analysis of magnetically tuned mass dampers, applied to the vibration damping of large structures of civil engineering interest. Two devices are analysed, for which both the frequency tuning ratio and the damping coefficient can be easily and finely calibrated. They are applied for the damping of the vibrations along two natural modes of a mock-up of a bridge under construction. An original analysis, based on the Maxwell receding image method, is developed for estimating the drag force arising inside the damping devices. It also takes into account self inductance effects, yielding a complex nonlinear dependence of the drag force on the velocity. The analysis highlights the range of velocities for which the drag force can be assumed of viscous type, and shows its dependence on the involved geometrical parameters of the dampers. The model outcomes are then compared to the corresponding experimental calibration curves. A dynamic model of the controlled structure equiped with the two damping devices is presented, and used for the development of original optimization expressions and for determining the corresponding maximum achievable damping. Finally, several experimental results are presented, concerning both the free and harmonically forced vibration damping of the bridge mock-up, and compared to the corresponding theoretical predictions. The experimental results reveal that the maximum theoretical damping performance can be achieved, when both the tuning frequencies and damping coefficients of each device are finely calibrated according to the optimization expressions [13], [44].

6.2.2. Particle filtering techniques for monitoring of structures

Participant: Laurent Mevel.

The focus of this paper is Bayesian modal parameter recursive estimation based on an interacting Kalman filter algorithm with decoupled distributions for frequency and damping. Interacting Kalman filter is a combination of two widely used Bayesian estimation methods: the particle filter and the Kalman filter. Some sensitivity analysis techniques are also proposed in order to deduce a recursive estimate of modal parameters from the estimates of the damping/stiffness coefficients [28].

6.2.3. Uncertainty quantification

Participants: Michael Doehler, Laurent Mevel.

For applications as Operational Modal Analysis (OMA) of vibrating structures, an output-only LTI system with state and measurement noise can be identified using subspace methods. While these identification techniques have been very suitable for the identification of such mechanical, aeronautical or civil structures, covariance expressions of the estimates of the system matrices are difficult to obtain and theoretical results from literature are hard to implement for output-only systems with unknown noise properties in practice. Moreover, the model order of the underlying system is generally unknown and due to noise and model errors, usual statistical criteria cannot be used. Instead, the system is estimated at multiple model orders and some GUI driven stabilization diagram containing the resulting modal parameters is used by the structural engineer. Then, the covariance of the estimates at these different model orders is an important information for the engineer, which, however, would be computationally expensive to obtain with the existing tools. Recently a fast multi-order version of the stochastic subspace identification approach has been proposed, which is based on the use of the QR decomposition of the observability matrix at the largest model order. In this paper, the corresponding covariance expressions for the system matrix estimates at multiple model orders are derived and successfully applied on real vibration data [36], [38].

6.2.4. Periodic systems

Participants: Ivan Guéguen, Laurent Mevel.

The modal analysis of a wind turbine has been generally handled with the assumption that this structure can be accurately modeled as linear time-invariant. Such assumption may be misleading for stability analysis, especially, with the current development of very large wind turbines with complex dynamic behavior (nonlinearity, aeroelastic coupling). Therefore in this paper, the inherent periodically time-varying dynamics of wind turbines (and for rotating systems, in general) is taken into account. Recently a subspace algorithm for modal analysis of rotating systems has been proposed. It is tested on a simulated and real data from a wind turbine [20], [41].

6.2.5. Identification of finite impulse response systems based on quantized output measurements – a quadratic programming-based method

Participant: Qinghua Zhang.

This work has been carried out in collaboration with Jiandong Wang (Peking University, China).

Quantized data are typically produced by the process of analog-to-digital conversion and have been widely studied in signal encoding and digital representation. In system identification, the processed data are usually collected after a quantization procedure, but the effect of quantization is often ignored. The study on system identification based on quantized data makes sense when the data are coded with few quantization levels, to the point that the effect of quantization becomes important. In this work we propose a quadratic programming (QP)-based method for identification of finite impulse response (FIR) dynamic systems from quantized or binary data. The main idea of the proposed method is to reformulate this identification problem, usually viewed as a nonlinear estimation problem with discontinuous nonlinearities, in the form of a standard QP problem, which is a convex optimization problem and can be solved efficiently. The complete input conditions ensuring the strict convexity of the QP problem are developed, and the consistency of the estimated parameters is established under the complete input conditions. The results of this study have been published in [27].

6.2.6. Wiener System Identification by Weighted Principal Component Analysis

Participant: Qinghua Zhang.

This work has been carried out in collaboration with Vincent Laurain (CRAN/CNRS/Université de Lorraine).

A Wiener system consists of two subsystems connected in series, with a linear dynamic subsystem preceding a static nonlinearity. In the field of control systems, the dynamics of a nonlinear system can often be linearized around its working point. Nevertheless, if its output sensor is affected by strongly nonlinear distortions, the linearization of the sensor characteristics may induce large modeling errors. In such situations, Wiener system model is more appropriate than fully linearized models. Wiener system identification is investigated in this work with a finite impulse response (FIR) model of the linear subsystem. Under the assumption of Gaussian input distribution, this work mainly aims at addressing a deficiency of the well-known correlation-based method for Wiener system identification: it fails when the nonlinearity of the Wiener system is an even function. This method is, in the considered Gaussian input case, equivalent to the best linear approximation (BLA) method, which exhibits the same deficiency. Our new method is based on a weighted principal component analysis (wPCA). Its consistency is proved for Wiener systems with either even or non even nonlinearities. Its computational cost is almost the same as that of a standard PCA.The results of this study have been presented at [51].

6.2.7. Industrial process for road buildings

Participant: Jean Dumoulin.

The increasing use of the baffled-rotary kiln equipment in many innovative materials processing industrial applications suggests examining the heat transfer phenomena in order to improve the multi-phase flow modeling tools. Their development and use will be relevant for tackling the current energy issues. The heat transfer models available for the rotary kiln in the literature are, for now, not enough efficient for the baffled-rotary kiln case. The present paper is aimed at suggesting a wall heat transfer correlation for the rotary kilns with the secondary inlet. The experimental thermal data acquired within large-scale rotary drum applied to the asphalt concrete materials production, are remained in order to give rise the new issues. These latter results are connected to a visualization campaign performed at the pilot-scale in order to assess the transversal distribution of the granular phase materials. Their analysis suggests a more appropriate physical modelling of the wall heat transfer path. It leads to transform the classical correlation of type $Nu = f(Re, Pr)$ in a new expression of type $Nu = f(Re, St)$ based on a new physical modeling inventory corresponding to the hot and cold fluxes flowing within the baffled-rotary kiln. Thus, the major modification is based on the introduction of the Stanton (St) number in the wall heat transfer correlation. This expression is found more convenient for the baffled-rotary kiln application. This new expression is validated by the comparison with the experimental Nusselt numbers calculated from the inner heat transfer measurements coefficient measured in the baffled-rotary kiln performed at large scale [21].

6.2.8. Industrial process for concrete structure reparation

Participant: Jean Dumoulin.

In civil engineering, reinforced concrete repair by CFRP is a strengthening technique that has proven successfully in the past. The present study is aimed at using thermoplastic CFRP sheets applied and glued under heat. In this research framework, active thermography is used to accomplish two roles: control of the operating temperature of the thermoplastic CFRP sheets during the installation process and evaluation of the bonding quality after welding. The pap er presents results obtained in laboratory with a dedicated test bench coupled with numerical simulations of the process [49].

6.2.9. Building energy management

Participants: Alexandre Nassiopoulos, Jordan Brouns.

Problems such as parameter identification for model calibration, optimal design or optimal energy management can all be formulated in a similar framework as problems consisting in finding the minimum of a cost function. The paper presents the software ReTrofiT that specifically treats this kind of problems applied to building energy performance models. ReTrofiT is first of all a simulation tool for evaluating building thermal behavior and computing energy consumptions. The novelty compared to state-of-the-art energy simulation software is that it also integrates a generic set of tools and algorithms to set up and solve optimization problems related to the building thermal model. The use of the adjoint model, that is intrinsically implemented in the code, constructs fast and efficient algorithms to solve linear, non linear, constrained or unconstrained problems addressing a wide range of applications [43].

6.3. damage detection for mechanical structures

6.3.1. Damage detection and localisation

Participants: Michael Doehler, Luciano Gallegos, Laurent Mevel.

The Stochastic Dynamic Damage Locating Vector approach is a vibration-based damage localization method based on a finite element model of a structure and output-only measurements in both reference and damaged states. A stress field is computed for loads in the null space of a surrogate of the change in the transfer matrix at the sensor positions for some values in the Laplace domain. Then, the damage location is related to positions where the stress is close to zero. Robustness of the localization information can be achieved by aggregating results at different values in the Laplace domain. So far, this approach and in particular the aggregation is deterministic and does not take the uncertainty in the stress estimates into account. In this paper, the damage localization method is extended with a statistical framework. The uncertainty in the output-only measurements is propagated to the stress estimates at different values of the Laplace variable and these estimates are aggregated based on statistical principles. The performance of the new statistical approach is demonstrated both in a numerical application and a lab experiment, showing a significant improvement of the robustness of the method due to the statistical evaluation of the localization information [22], [37].

6.3.2. An Innovations Approach to Fault Diagnosis in Linear Time-Varying Descriptor Systems

Participant: Qinghua Zhang.

This work has been carried out in collaboration with Abdouramane Moussa-Ali (LSIS/CNRS/Université de Toulon).

Many modern engineering systems can be modeled by explicit ordinary differential equations (ODE) in state-space form. Such state-space equations have a long-term mathematical history, and a large number of analytical and numerical tools have been developed for their study. Nevertheless, some systems cannot be described by such explicit state-space models, but described by *implicit* differential equations, known as differential-algebraic equations (DAE). After linearization along a trajectory and discretization in time, a nonlinear DAE system is approximately described by *implicit* discrete time state-space equations, known as *descriptor system equations*. In this work, fault diagnosis is studied for time varying descriptor systems. The Kalman filter for descriptor systems is first revisited by completing existing results about its properties that are essential for the purpose of fault diagnosis. Based on the analysis of the effects of the considered actuator and sensor faults on the innovation of the descriptor system Kalman filter, it is shown that the considered fault diagnosis problem in time varying descriptor systems is equivalent to a classical linear regression problem formulated by appropriately filtering the input-output data. Following this result, algorithms for fault diagnosis through maximum likelihood estimation are then developed. The results of this study have been presented at [42].

6.3.3. Statistical detection and isolation of additive faults in linear time-varying systems

Participant: Qinghua Zhang.

This work has been carried out in collaboration with Michèle Basseville (IRISA/CNRS).

Model-based approaches to fault detection and isolation (FDI) have been mostly studied in the literature for linear time invariant (LTI) systems. In practice, quite often time-varying and/or nonlinear properties of the monitored system cannot be neglected. One of the possible approachs to dealing with nonlinear systems is based on the linearization along the actual or nominal trajectory of the monitored system. Such a linearization generally leads to linear time-varying (LTV) systems, whereas the more basic LTI approximation is usually related to the linearization around a single working point. It is thus clear that methods for FDI in LTV systems are much more powerful than their LTI counterparts. In the present work, we address the FDI problem for LTV systems subject to parametric additive faults. The proposed approach is statistical, by combining a generalized likelihood ratio (GLR) test with the Kalman filter that cancels out the dynamics of the faults effects in the considered LTV systems. With this approach, it is possible to perform fault isolation when the number of sensors is smaller than the number of assumed faults, under an appropriate assumption about the excitation of the system. The results of this study have been published in [29].

6.3.4. Robust subspace damage detection

Participants: Michael Doehler, Laurent Mevel.

In the last ten years, monitoring the integrity of the civil infrastructure has been an active research topic, including in connected areas as automatic control. It is common practice to perform damage detection by detecting changes in the modal parameters between a reference state and the current (possibly damaged) state from measured vibration data. Subspace methods enjoy some popularity in structural engineering, where large model orders have to be considered. In the context of detecting changes in the structural properties and the modal parameters linked to them, a subspace-based fault detection residual has been recently proposed and applied successfully, where the estimation of the modal parameters in the possibly damaged state is avoided. However, most works assume that the unmeasured ambient excitation properties during measurements of the structure in the reference and possibly damaged condition stay constant, which is hardly satisfied by any application. This paper addresses the problem of robustness of such fault detection methods. It is explained why current algorithms from literature fail when the excitation covariance changes and how they can be modified. Then, an efficient and fast subspace-based damage detection test is derived that is robust to changes in the excitation covariance but also to numerical instabilities that can arise easily in the computations. Three numerical applications show the efficiency of the new approach to better detect and separate different levels of damage even using a relatively low sample length [18], [35], [17].

6.3.5. Sensor placement

Participant: Michael Doehler.

Deciding on the position of sensors by optimizing the utility of the monitoring system over a structure lifetime is typically forbidden by computational cost. Sensor placement strategies are, instead, usually formulated for a pre-selected number of sensors and are based on cost functions that can be evaluated for any arrangement without the need for simulations. This paper examines the performance of two such schemes, the first one is derived directly from a technique that detects damage from the shift of a chi-square distribution from central to non-central and takes the optimal arrangement as the one that maximizes the sensitivity of the non-centrality to all parameter changes of equal norm. The second scheme selects the sensor arrangement as that which maximizes a weighted version of the norm of the sensitivity of the covariance of the output to all feasible changes in system parameters. The performance of the two schemes is tested in simulations [32].

6.3.6. Reflectometry for external post-tensioned cable monitoring

Participant: Qinghua Zhang.

This work has been carried out in collaboration with IFSTTAR, EDF, ENS Cachan and Andra.

Nowadays a considerable number of bridges is reaching an age when repairs become necessary. In some bridges, external post-tension cables are placed in ducts within which the residual internal space is imperfectly filled with a fluid cement grout. Detecting the defaults of filling is visually impossible from the outside. Among non-destructive detection techniques proposed for cable health monitoring, reflectometry techniques offer remarkable advantages in that they can monitor cables in concrete deviator (embedded in concrete) and they do not require human intervention inside the bridge. In this work, the application of reflectometry techniques to cable health monitoring has been investigated via numerical simulations and laboratory experiments. The results of this study have been presented at [53].

6.3.7. Efficient Computation of Minmax Tests for Fault Isolation and Their Application to Structural Damage Localization

Participants: Michael Doehler, Laurent Mevel.

Fault detection and isolation can be handled by many different approaches. This paper builds upon a hypothesis test that checks whether the mean of a Gaussian random vector has become non-zero in the faulty state, based on a chi2 test. For fault isolation, it has to be decided which components in the parameter set of the Gaussian vector have changed, which is done by variants of the chi2 hypothesis test using the so-called sensitivity and minmax approaches. While only the sensitivity of the tested parameter component is taken into account in the sensitivity approach, the sensitivities of all parameters are used in the minmax approach, leading to better statistical properties at the expense of an increased computational burden. The computation of the respective test variable in the minmax test is cumbersome and may be ill-conditioned especially for large parameter sets, asking hence for a careful numerical evaluation. Furthermore, the fault isolation procedure requires the repetitive calculation of the test variable for each of the parameter components that are tested for a change, which may be a significant computational burden. In this paper, dealing with the minmax problem, we propose a new efficient computation for the test variables, which is based on a simultaneous QR decomposition for all parameters. Based on this scheme, we propose an efficient test computation for a large parameter set, leading to a decrease in the numerical complexity by one order of magnitude in the total number of parameters. Finally, we show how the minmax test is useful for structural damage localization, where an asymptotically Gaussian residual vector is computed from output-only vibration data of a mechanical or a civil structure [39].

6.3.8. Inverse problems in damage detection

Participant: Dominique Siegert.

Reinforced concrete beams are widely employed in civil engineering structures. To reduce the maintenance financial cost, structure damages have to be detected early. To this end, one needs robust monitoring techniques. The paper deals with the identification of mechanical parameters, useful for Structural Health Monitoring, in a 2D beam using inverse modeling technique. The optimal control theory is employed. As an example, we aim to identify a reduction of the steel bar cross-section and a decrease of the concrete Young modulus in damaged areas. In our strategy, the beam is instrumented with strain sensors, and a known dynamic load is applied. In the inverse technique, two space discretizations are considered: a fine dicretization to solve the structural dynamic problem and a coarse discretization for the beam parameter identification. To get the beam parameters, we minimize a classical data misfit functional using a gradient-like algorithm. A low-cost computation of the functional gradient is performed using the adjoint equation. The inverse problem is solved in a general way using engineer numerical tools: Python scripts and the free finite element software Code Aster. First results show that a local reduction of the steel bar cross-section and a local decrease of concrete Young modulus can be detected using this inverse technique [25].

6.3.9. NDT by active thermography coupled with infrared shearography

Participant: Jean Dumoulin.

As infrastructures are aging, the evaluation of their health is becoming crucial. To do so, numerous Non Destructive Testing (NDT) methods are available. Among them, thermal shearography and active infrared thermography represent two full field and contactless methods for surface inspection. The synchronized use of both methods presents multiples advantages. Most importantly, both NDT are based on different material properties. Thermography depend on the thermal properties and shearography on the mechanical properties. The cross-correlation of both methods result in a more accurate and exact detection of the defects. For real site application, the simultaneous use of both methods is simplified due to the fact that the excitation method (thermal) is the same. Active infrared thermography is the measure of the temperature by an infrared camera of a surface subjected to heat flux. Observation of the variation of temperature in function of time reveal the presence of defects. On the other hand, shearography is a measure of out-of-plane surface displacement. This displacement is caused by the application of a strain on the surface which (in our case) take the form of a temperature gradient inducing a thermal stress [56], [47], [48].

6.4. Long term monitoring of civil engineering structure

6.4.1. ICT based software for thermal field long term monitoring of civil engineering structures

Participants: Antoine Crinière, Jean Dumoulin.

Aging of transport infrastructures combined with traffic and climatic solicitations contribute to the reduction of their performances. To address and quantify the resilience of civil engineering structure, investigations on robust, fast and efficient methods are required. Among research works carried out at IFSTTAR, methods for long term monitoring face an increasing demand. Such works take benefits of this last decade technological progresses in ICT domain. A multi-sensing techniques system, able to date and synchronize measurements carried out by infrared thermography coupled with various measurements data (i.e. weather parameters), have been designed, developed and implemented on real site. This smart sensor called IrLaw/SENSORBOX has been upgraded in order to reach full autonomy and its able to monitor over years civil engineering structures [55], [15], [34].

6.4.2. Long term structural health monitoring architecture

Participant: Jean Dumoulin.

This work gives a brief description of the main activities and outcomes of the Integrated System for Transport Infrastructures surveillance and Monitoring by Electromagnetic Sensing (ISTIMES – www.istimes.eu) project, which was concerned with the development and implementation of a system able to couple the capabilities of long-term monitoring and quick damage assessment of the critical transport infrastructures. This was performed thanks to the integrated use of the novel and state of art concepts of Earth observation, ground-based sensing techniques and ICT architecture [45], [46].

6.5. Material characterization

6.5.1. Quantitative non destructive testing in civil engineering

Participants: Jordan Brouns, Antoine Crinière, Jean Dumoulin, Alexandre Nassiopoulos.

By the aging of civil engineering structures a crucial need of reparation or reinforcement appeared through years. This can be done using bonded CFRP plate to assure the mechanical behavior of the structure. This type of reparation need diagnosis to insure the reliability of the reparation procedure. This part focus on the development of 1D to 3D method to asses the quantitative non destructive testing of a repaired structure thanks to active thermography (see [14] and [52]).

6.5.2. Thermo-physical characterization for civil engineering application

Participant: Jean Dumoulin.

This papers presents the development of a new device for the determination of thermal conductivity and diffusivity of anisotropic composite plates. The excitation signal is provided through a thermoelectric cooler and does not require any optical source like a laser source for instance. Infrared thermography is used to follow apparent surface temperature evolution with time. Experiments were carried out two composite sample systems (with two different fiber orientations). Result analysis is presented and discussed [40].

6.5.3. Emissivity characterization for civil engineering applications

Participant: Jean Dumoulin.

The knowledge of the infrared emissivity of materials used in buildings and civil engineering structures is useful for two specific approaches. First, quantitative diagnosis of buildings or civil engineering infrastructures by infrared thermography requires emissivity values in the spectral bandwidth of the camera used for measurements, in order to obtain accurate surface temperatures; for instance, emissivity in the band III domain is required when using cameras with uncooled detectors (such as micro-bolometer arrays). Second, setting up accurate thermal balances by numerical modeling requires the total emissivity value for a large wavelength domain; this is, for instance, the case for computing the road surface temperature to predict ice occurrence. Furthermore, periodical surveys of emissivity variations due to aging or soiling of surfaces could be useful in many situations such as thermal mapping of roads or building insulation diagnosis. The use of portable emissivity measurement devices is required for that purpose. A device using an indirect measurement method was previously developed in our lab; the method uses measurement of the reflectivity from a modulated IR source and requires calibration with a highly reflective surface. However, that device uses a low-frequency, thermal modulation well adapted to laboratory measurements but unfit for fast and in situ measurements. Therefore, a new, portable system which retains the principle of an indirect measurement but uses a faster-frequency, mechanical modulation more appropriate to outdoor measurements was developed. Both devices allow measurements in the broad (1μm to 40μm) and narrow (8μm to 40μm) bands. Experiments were performed on a large number of materials commonly used in buildings and civil engineering structures.The final objective of this work is to build a database of emissivity of these materials. A comparison of laboratory and on-site measurements of emissivity values obtained in both spectral bands will be presented along with an estimation and an analysis of measurement uncertainties [23].

6.6. Vision under environmental conditions

6.6.1. Infrared Imaging under environmental conditions

Participant: Jean Dumoulin.

An infrared system has been developed to monitor transport infrastructures in a standalone configuration. It is based on low cost infrared thermal cameras linked with a calculation unit in order to produce a corrected thermal map of the surveyed structure at a selected time step. With the inline version, the data collected feed simplified radiative models running a GPU. With the offline version, the thermal map can be corrected when data are collected under different atmospheric conditions up to foggy night conditions. A model for radiative transmission prediction is proposed and limitations are addressed. Furthermore, the results obtained by image and signal processing methods with data acquired on the transport infrastructure opened to traffic are presented. Finally, conclusions and perspectives for new implementation and new functionalities are presented and discussed [16].

6.6.2. Long term thermal monitoring by uncooled infrared camera

Participant: Jean Dumoulin.

Being able to perform easily non-invasive diagnostics for surveillance and monitoring of critical transport infrastructures is a major preoccupation of many technical offices. Among all the existing electromagneticmethods, long term thermal monitoring by uncooled infrared camera is a promising technique due to its dissemination potential according to its low cost on the market. Nevertheless, Knowledge of environmental parameters during measurement in outdoor applications is required to carry out accurate measurement corrections induced by atmospheric effects at ground level. Particularly considering atmospheric effects and measurements in foggy conditions close as possible to those that can be encountered around transport infrastructures, both in visible and infrared spectra. In the present study, atmospheric effects are first addressed by using data base available in literature and modelling. Atmospheric attenuation by particles depends greatly of aerosols density, but when relative humidity increases, water vapor condenses onto the particulates suspended in the atmosphere. This condensed water increases the size of the aerosols and changes their composition and their effective refractive index. The resulting effect of the aerosols on the absorption and scattering of radiation will correspondingly be modified [54].

6.6.3. Handling of fog conditions by infrared cameras

Participant: Jean Dumoulin.

Fog conditions are the cause of severe car accidents in western countries because of the poor induced visibility. Its forecast and intensity are still very difficult to predict by weather services. Infrared cameras allow to detect and to identify objects in fog while visibility is too low for eye detection. Over the past years, the implementation of cost effective infrared cameras on some vehicles has enabled such detection. On the other hand pattern recognition algorithms based on Canny filters and Hough transformation are a common tool applied to images. Based on these facts, a joint research program between IFSTTAR and Cerema has been developed to study the benefit of infrared images obtained in a fog tunnel during its natural dissipation. Pattern recognition algorithms have been applied, specifically on road signs which shape is usually associated to a specific meaning (circular for a speed limit, triangle for an alert, ...). It has been shown that road signs were detected early enough in images, with respect to images in the visible spectrum, to trigger useful alerts for Advanced Driver Assistance Systems [33].

IBIS Project-Team

5. New Results

5.1. Highlights of the Year

A paper based on the PhD thesis of Diana Stefan was accepted for *PLoS Computational Biology* this year [7].

5.2. Analysis of gene regulatory networks by means of piecewise-linear (PL) models

GENETIC NETWORK ANALYZER (GNA) is a tool for the qualitative modeling and simulation of the dynamics of gene regulatory networks by means of PLDE models, as described in Section 4.1 . GNA has been integrated with the other bioinformatics tools distributed by Genostar (http://www.genostar.com/). Version 8.7 of GNA was released by IBIS and Genostar this year. This version is an update of version 8.0, deposited at the Agence pour la Protection des Programmes (APP). Some bugs have been corrected in the new version and the program has been adapted to the latest versions of Java and the software platform of Genostar. Version 8.7 supports the SBML standard and is also capable of exporting its models to the newly-developed standard for qualitative models, SBML Qual. This standard has been elaborated by the community of developers of logical and related modeling tools (CoLoMoTo), in which the GNA developers participate.

The predictions obtained with the help of GNA are purely qualitative, describing the dynamics of the network by means of a state transition graph. While a qualitative analysis is appropriate for certain problems, the absence of precise quantitative predictions may not be desirable in others, such as the analysis of a limit cycle or the design of a controller for a synthetic network. The quantitative study of PLDE models of gene regulatory networks is hindered by the fact that the step functions describing the logic of regulatory interactions lead to discontinuities in the right-hand side of the PLDE models (Section 3.1). This has motivated extensions of the PLDE models based on differential inclusions and Filippov solutions. As of now, no numerical simulation tool for the simulation of these Filippov extensions is available.

In collaboration with the BIPOP project-team, we have shown how tools developed for the simulation of non-smooth mechanical, electrical and control systems can be adapted for this purpose, in a paper published in *Physica D* [1] and presented at the 21st International Symposium on Mathematical Theory of Networks and Systems (MTNS 2014) [12]. We have presented a method for the numerical analysis of one proposed extension, called Aizerman–Pyatnitskii (AP)-extension, by reformulating the PLDE models as mixed complementarity systems (MCSs). This allows the application of powerful methods developed for this class of nonsmooth dynamical systems, in particular those implemented in the SICONOS platform developed by BIPOP. We have also shown that under a set of reasonable biological assumptions, putting constraints on the right-hand side of the PLDE models, AP-extensions and classical Filippov extensions are equivalent. This means that the proposed numerical method is valid for a range of different solution concepts. We have illustrated the practical interest of our approach through the numerical analysis of three well-known networks developed in the field of synthetic biology.

5.3. Inference of bacterial regulatory networks from reporter gene data

The use of fluorescent and luminescent reporter genes allows real-time monitoring of gene expression, both at the level of individual cells and cell populations (Section 3.2). In order to fully exploit this technology, we need methods to rapidly construct reporter genes, both on plasmids and on the chromosome, mathematical models to infer biologically relevant quantities from the primary data, and computer tools to achieve this in an efficient and user-friendly manner. For instance, in a typical microplate experiment, 96 cultures are followed in parallel, over several hours, resulting in 10,000-100,000 measurements of absorbance and fluorescence and luminescence intensities. Over the past few years, we put into place an experimental platform and data analysis software, notably the WELLREADER program (Section 4.2), to allow biologists to make the most out of the information contained in reporter gene expression data.

Valentin Zulkower, in the framework of his PhD thesis, has developed novel methods for the analysis of reporter gene data, based on the use of regularized linear inversion. This allows a range of estimation problems in the analysis of reporter gene data, notably the inference of growth rate, promoter activity, and protein concentration profiles, to be solved in a mathematically sound and practical manner. We have evaluated the validity of the approach using *in-silico* simulation studies, and observed that the methods are more robust and less biased than indirect approaches usually encountered in the experimental literature based on smoothing and subsequent processing of the primary data, like in WELLREADER. We have applied the methods to the analysis of fluorescent reporter gene data acquired in kinetic experiments with *Escherichia coli*. The methods were shown capable of reliably reconstructing time-course profiles of growth rate, promoter activity, and protein concentration from weak and noisy signals at low population volumes. Moreover, they captured critical features of those profiles, notably rapid changes in gene expression during growth transitions. The linear inversion methods have been implemented in the Python package WELLFARE, and integrated by Michel Page in the web application WELLINVERTER (Section 4.2). This work was submitted for publication early 2015.

The above tools have been used in a series of studies directed at the experimental mapping of gene regulatory networks in *E. coli*. A first example is a study, led by Stéphan Lacour in collaboration with Akira Ishihama and Hiroshi Ogasawara in Japan, on the lifestyle adaptation of *E. coli*. The study concerns the switch between swimming motility and biofilm formation in response to changes in environmental growth conditions. The stationary phase sigma factor RpoS is an important regulator of this switch since it stimulates adhesion and represses flagellar biosynthesis. By measuring the dynamics of gene expression, we show that RpoS inhibits the transcription of the flagellar sigma factor, FliA, in exponential growth phase. RpoS also partially controls the expression of CsgD and CpxR, two transcription factors important for bacterial adhesion. We have demonstrated that these two regulators repress the transcription of *fliA*, *flgM* and *tar*, and that this regulation is dependent on the growth medium. CsgD binds to the flgM and fliA promoters around their -10 promoter element, strongly suggesting direct repression. The results show that CsgD and CpxR also affect the expression of other known modulators of cell motility. An updated structure of the regulatory network controlling the choice between adhesion and motility was proposed in the paper based on this work, published in the *Journal of Bacteriology* [2]. Stéphan Lacour also reviewed this and other work on RpoS in a publication in *Environmental Microbiology Reports* [4].

A second example derives from the PhD thesis of Diana Stefan. Although from a biological point of view the motility network of *E. coli* is also central in this work, its main thrust lies in clarifying and solving methodological issues in the automated inference of quantitative models of gene regulatory networks from time-series gene expression data, also called reverse engineering in the bioinformatics literature. The application of existing reverse engineering methods is commonly based on implicit assumptions on the biological processes under study. First, the measurements of mRNA abundance obtained in transcriptomics experiments are taken to be representative of protein concentrations. Second, the observed changes in gene expression are assumed to be solely due to transcription factors and other specific regulators, while changes in the activity of the gene expression machinery and other global physiological effects are neglected. While convenient in practice, these assumptions are often not valid and bias the reverse engineering process. In her PhD thesis, Diana Stefan systematically investigated, using a combination of models and experiments, the importance of this bias and possible corrections. She measured with the help of fluorescent reporter genes the activity of genes involved in the FliA-FlgM module of the *E. coli* motility network. From these data, protein concentrations and global physiological effects were estimated by means of kinetic models of gene expression. The results indicate that correcting for the bias of commonly-made assumptions improves the quality of the models inferred from the data. Moreover, it was shown by simulation that these improvements are expected to be even stronger for systems in which protein concentrations have longer half-lives and the activity of the gene expression machinery varies more strongly across conditions than in the FliA-FlgM module. The approach proposed in this study is broadly applicable when using time-series transcriptome data to learn about the structure and dynamics of regulatory networks. The paper describing the work was published in *PLoS Computational Biology* [7].

5.4. Models of carbon metabolism in bacteria

All free-living bacteria have to adapt to a changing environment. Specific regulatory systems respond to particular stresses, but the most common decision bacteria have to make is the choice between alternative carbon sources, each sustaining a specific, maximal growth rate. Many bacteria have evolved a strategy that consists in utilizing carbon sources sequentially, in general favouring carbon sources that sustain a higher growth rate. As long as a preferred carbon source is present in sufficient amounts, the synthesis of enzymes necessary for the uptake and metabolism of less favourable carbon sources is repressed. This phenomenon is called Carbon Catabolite Repression (CCR) and the most salient manifestation of this regulatory choice is diauxic growth, a phenomenon discovered by Jacques Monod more than 70 years ago. Although this system is one of the paradigms of the regulation of gene expression in bacteria, the underlying mechanisms remain controversial. Carbon catabolite repression involves the coordination of different subsystems of the cell - responsible for the uptake of carbon sources, their breakdown for the production of energy and precursors, and the conversion of the latter to biomass.

The complexity of this integrated system, with regulatory mechanisms cutting across metabolism, gene expression, signaling and subject to global physical and physiological constraints, has motivated important modeling efforts over the past four decades, especially in the enterobacterium *Escherichia coli*. Different hypotheses concerning the dynamic functioning of the system have been explored by a variety of modeling approaches. In an article in *Trends in Microbiology* [3], which was initiated during the sabbatical of Andreas Kremling in Grenoble in 2013, we have reviewed these studies and summarized their contributions to the quantitative understanding of carbon catabolite repression, focusing on diauxic growth in E. coli. Moreover, we have proposed a highly simplified representation of diauxic growth that makes it possible to bring out the salient features of the models proposed in the literature and confront and compare the explanations they provide.

A bottleneck in the development of dynamic and quantitatively predictive models of bacterial metabolism, explicitly accounting for the different regulatory mechanisms on the molecular level, is information on the kinetic parameters describing the enzymatic reactions and other molecular interactions. One particularly important piece of information is knowledge of enzyme concentrations. Recent technological advances in quantitative proteomics have made mass spectrometry-based quantitative assays an interesting alternative to more traditional immuno- affinity based approaches for quantifying enzyme concentrations. In particular, these advances have improved specificity and multiplexing capabilities. In a study carried out at CEA Grenoble, a quantification workflow to analyze enzymes involved in central metabolism in *E. coli* was developed. This workflow combined full-length isotopically labeled standards with selected reaction monitoring analysis. The workflow was used to accurately quantify 22 enzymes involved in *E. coli* central metabolism in a wild-type reference strain and two derived strains, optimized for higher NADPH production. Delphine Ropers and Hidde de Jong participated in the analysis of these data. In combination with measurements of metabolic fluxes, we showed that proteomics data can be used to assess different levels of regulation, in particular enzyme abundance and catalytic rate. This is key to the development of predictive kinetic models, but also provides information that can be used for strain design in biotechnology. An article based on this work was published in *Molecular and Cellular Proteomics* [8].

Other ongoing work on the analysis of bacterial metabolism is carried out by Delphine Ropers in collaboration with Inra/INSA in Toulouse, in the framework of the PhD thesis of Manon Morin, supported by a Contrat Jeune Scientifique Inra-Inria. In their respective PhD theses, Stéphane Pinhal and Valentin Zulkower also study specific aspects of carbon metabolism, using both models and experimental data.

5.5. Stochastic modeling and identification of gene regulatory networks in bacteria

At the single-cell level, the processes that govern single-cell dynamics in general and gene expression in particular are better described by stochastic models. Modern techniques for the real-time monitoring of gene expression in single cells enable one to apply stochastic modelling to study the origins and consequences of random noise in response to various environmental stresses, and the emergence of phenotypic variability. The potential impact of single-cell stochastic analysis and modelling ranges from a better comprehension of the biochemical regulatory mechanisms underlying cellular phenotypes to the development of new strategies for the (computer assisted or genetically synthesized) control of cell populations and even of single cells.

Work in IBIS on the probabilistic gene expression and interaction dynamics at the level of individual cells is addressed in terms of identification of intrinsic noise models from population snapshot data, on the one hand, and the inference of models focusing on cellular variability within isogenic populations from individual cell fluorescence microscopy gene expression profiles, on the other hand. Along with modelling and inference comes analysis of the inferred models in various respects, notably in terms of single-cell state estimation and control. Other problems related with single-cell modelling and extracellular variability are considered in high-eukariotic cells through external collaborations.

In the context of yeast cell response to osmotic shocks, in collaboration with the CONTRAINTES project-team, and colleagues from Université Paris Descartes and University of Pavia (Italy), Eugenio Cinquemani has investigated the use of mixed effects-modelling and identification techniques to characterize individual cell dynamics in isogenic cell populations. Mixed-effects models are hierarchical models where parametric response profiles of individuals is subject to inter-individual parameter variability following a common population distribution. Starting from identification approaches in pharmacokinetics, we have developed and applied inference methods to the context of microfluidics data, with focus on the budding yeast response to osmotic shocks. First results presented at conference in 2013 have been taken further, both in terms of mathematical analysis of the models developed and in terms of biological interpretation. Model identification and validation were performed together with Andres Gonzales, PhD student at the University of Pavia, who has visited IBIS for six months in 2014. A journal publication is currently being prepared for publication.

In a second line of work, starting from the models inferred in the above collaboration, the problem of real-time state estimation and control of single yeast cells has been considered. Together with the BIOCORE project-team, we have put in place algorithms for state estimation in presence of hybrid random switching and continuous dynamics, and integrated them with a feedback control approach developed by collaborators at TU Delft (the Netherlands). The whole monitoring, estimation and control chain has been deployed and applied *in silico* to the stochastic control of osmosensitive genes in single yeast cells. Methods and results have been presented at the 12th international conference on Computational Methods for Systems Biology (CMSB 2014), whose proceedings have been published as a volume of the LNCS series [14]. It is shown in particular that stochastic model-based estimation and control outperforms existing methods of single-cell control based on deterministic approximations.

Additional work on identification and estimation of hidden states for intrinsic noise models of gene expression/regulation in single bacterial cells, started with reference to arabinose uptake dynamics but also applicable to other regulatory networks in *E. coli*, is being developed. In parallel, collaboration of Eugenio Cinquemani with Marianna Rapsomaniki, PhD student affiliated with the University of Patras (Greece) and ETH Zürich (Switzerland), has been devoted to the analysis of data from Fluorescence Recovery After Photobleaching (FRAP) experiments and the inference of kinetic parameters of protein dynamics in single high-eukariotic cells. As an alternative to current approximate analytical methods, we have explored inference methods based on simulation of biological processes in realistic environments at a particle level. We introduced and demonstrated a new method for the inference of kinetic parameters of protein dynamics, where a limited number of *in-silico* FRAP experiments is used to construct a mapping from FRAP recovery curves to the parameters sought. Parameter estimates from experimental data are then computed by applying the mapping to the observed recovery curves, at virtually no additional price for any number of experiments, along with the application of a bootstrap procedure for determining identifiability of the parameters and confidence intervals for their estimates. After validation on synthetic data, the method was successfully applied to the analysis of the nuclear proteins Cdt1, PCNA and GFPnls in mammalian cells, also shedding light on cell-to-cell variability of the protein kinetics. Method and results have recently been published in *Bioinformatics* [6].

5.6. Growth control in bacteria and biotechnological applications

A bacterial cell adapts its growth rate and the level of gene expression required to sustain growth to the environment, notably to the availability of nutrients providing the molecular building blocks and the energy required for growth. This adaptive response involves the global physiological state of the cell, in particular the activity of the gene expression machinery, and DNA-binding transcription factors and other specific regulators. While many studies have focused on networks of transcription factors, the analysis of the relative contributions of both transcription factors and global effects of the physiological state has received relatively little attention thus far. There is a huge literature on the molecular mechanisms coupling the activity of the gene expression machinery to changes in the nutritional quality of the environment, but a quantitative and dynamic picture of this very complicated regulatory system is still missing. Delphine Ropers and Edith Grac as well as Nils Giordano are developing models to achieve this, from bottom-up and top-down perspectives, respectively.

The quantitative models adopting the bottom-up pespective describe the molecular mechanisms controlling the activity of the gene expression machinery. The calibration and analysis of these models is made difficult by their complexity, the nonidentifiability of many parameter values, and the heterogeneity of experimental data sources. To overcome these difficulties, Delphine Ropers and Edith Grac are developing model ensembles with the same structure but different parameter values that are consistent with the experimental data. In collaboration with Jean-Luc Gouzé and Ismail Belgacem from the BIOCORE project-team at Inria Sophia-Antipolis-Méditerranée, they have analysed the dynamical behavior of a central module of these models, which controls the cellular concentration of the RNA polymerase, the key player of the transcriptional machinery. By means of model reduction approaches and monotone system theory, they have analyzed the equilibria of the system and their stability, which they could relate to biological observations on *E. coli*. This work has been published in the proceedings of the 21st International Symposium on Mathematical Theory of Networks and Systems (MTNS 2014) [9] and the 53rd IEEE Conference on Decision and Control (CDC 2014) [10]. A journal article is in preparation.

In the context of the PhD thesis of former IBIS member Jérôme Izard, we have studied the relation between the gene expression machinery, the global physiology of the cell, and the growth rate from a different perspective. Our aim was to change the mechanisms regulating the activity of the gene expression machinery in such a way so as to be able to externally control the growth rate of the cell. More precisely, we have engineered an *E. coli* strain in which the transcription of an essential component of the global gene expression machinery is under the tight control of an inducible promoter. By adjusting the inducer concentration in the medium we can adjust the activity of the gene expression machinery and thereby reversibly switch the growth rate of the bacterium between zero and the maximal growth rate. Our modified *E. coli* strain, described in a paper prepared for submission, opens new perspectives for studying the mechanisms of growth control as well as for developing biotechnological applications, the subject of the post-doctoral fellowship of Cindy Gomez Balderas-Barillot. We have submitted a patent proposing such applications, which underlies the technology transfer activities undertaken in the recently-started Reset project (Section 7.1).

IMAGINE Project-Team

6. New Results

6.1. Highlights of the Year

- Vector Graphics Complexes, an new structure for 2D illustration developed in collaboration with UBC, resulted into a publication at ACM SIGGRAPH [4]. This superset of multi-layers graphics and of planar maps, enable intuitive design and deformation of 2D illustrations thanks to the separation of geometry from topology.
- Our work on elastic implicit skinning, a collaboration with U. Toulouse, Victoria University, and Inria Bordeaux was accepted at ACM SIGGRAPH Asia [16]. Thanks to robust iso-surface tracking, this method captures dynamic skin siding effects and can be used with extreme bending angles.

6.2. User-centered Models for Shapes

- **Scientist in charge**: Stefanie Hahmann.
- **Other permanent researchers**: Marie-Paule Cani, Jean-Claude Léon.

Our goal, is to develop responsive shape models, i.e. 3D models that respond in the expected way under any user action, by maintaining specific application-dependent constraints (such as a volumetric objects keeping their volume when bent, or cloth-like surfaces remaining developable during deformation, etc). We are extending this approach to composite objects made of distributions and/or combination of sub-shapes of various dimensions.

6.2.1. Implicit modeling

Participants: Antoine Bégault, Marie-Paule Cani, Michael Gleicher, Cédric Zanni.

Figure 4. Illustration from [17] showing some results of our N-ary implicit blends.

Our insight towards 3D shapes that respond in an intuitive way during both design and animation is to develop representations that clearly separate changes of structure - namely, the morphology of the shape - from changes of posture (its current 3D isometric embedding). Using skeletons is an excellent way to do so for 3D solids: the structure of a shape is represented by the topology of the skeleton, the length of its components and the shape thickness around it, while the shape posture is defined by the embedding of the skeleton in 3D space. Implicit surfaces (iso-surfaces of scalar fields) are the best mathematical model so far for generating 3D shapes from skeletons. However, a number of long standing problems - blending at distance that makes topology unpredictable, bulges at junctions, blurring of details - reduced the interest for this representation for many years. We addressed several of these issues in the last few years. Our most recent contribution is a method for enabling topology control in the case n-ary implicit blends [17]. Shapes are modeled using scale-invariant integral primitives (SCALIS) along skeletons, and blend with a plus. We use field warping to avoid unwanted blending and provide a unique control (based for instance on the angle) on the way skeleton-based primitives are allowed to blend. See Figure 4

6.2.2. Towards responsive assemblies

Participants: Stefanie Hahmann, Jean-Claude Léon, Aarohi Singh Johal.

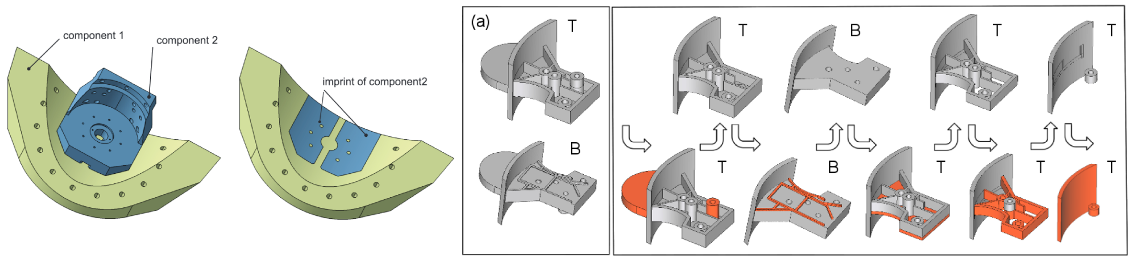


Figure 5. Illustration from [8] (Left) showing the automatic extraction of geometric interfaces, and from [2] (Right) showing the automatic generation of a construction graph.

We chose to focus on man-made objects to tackle the topic of shape assemblies, since CAD models of virtual industrial prototypes provide an excellent, real-size test-bed for our methods. Moreover, this is perfectly fitting the demand from industrial partners such as Airbus group and EDF.

Assemblies representing products are most often reduced to a collection of independent CAD models representing each component. The designation of each component and information about its function are often missing. As a result, geometric interfaces between components are unknown. These interfaces are particularly useful for structural mechanics to be able to quickly generate a Finite Element model of the assembly. This is especially critical when the latter gets very complex. [8] addresses the problem of automatically generating a class of geometric interfaces for very complex assemblies. GPU-based algorithms have proved suitable to obtain reliable results on CAD models.

Precisely determining interfaces between components is also a first requirement to enrich geometric models with functional information, since a subset of functions derives from interfaces between components. Based on both geometric interfaces and on a new concept of conventional interfaces, we proposed a series of approaches [13], [3] that make use of qualitative and ontology-based reasoning to connect CAD components and their geometric interfaces to functions or to functional designations of components: this results into an intrinsic identifier of a component in an assembly that connects it to its function.

To efficiently process assemblies of components, shape analysis [40] is particularly useful to generate the dimensionally reduced models needed for structural mechanics. [2] shows that analyzing a B-Rep CAD model to derive a construction graph, i.e. a set of construction trees, can be a robust basis to generate dimensionally reduced models.

Lastly, we extended shape analysis methods to detect some sets of symmetries [9]. Recovering this knowledge and embedding it into a model is the first step towards functionality-preserving deformations of complex man-made prototypes.

6.2.3. Parametric shapes

Participants: Stefanie Hahmann, Léo Allemand-Giorgis, Tibor Stanko.

Figure 6. Illustration from [38](Left) showing our results on monotonic interpolation, and from [1](Right) with our G1 interpolation surfaces for quad meshes.

We are developing new smooth parametric surface models defined on irregular quad meshes. They are in fact a powerful alternative to subdivision surface and singularly parameterized tensor product surfaces since they combine the advantages of both, the arbitrary topology of quad meshes and the smoothness of the tensor product patches. In collaboration with G.-P. Bonneau (Maverick team) several parametric triangular surface models for arbitrary topologies have been developed in the past. A new surface spline model has been published [1] and presented at GMP'14. It solves the problem of defining a G^1-continuous surface interpolating the vertices of an arbitrary quad mesh with low degree polynomial tensor product patches. It further aims to produce shapes of very high visual quality while reducing the number of control points, see Figure 6(right).

Another contribution concerns the modeling and smoothing of shapes using the Morse-Smale complex. The Morse-Smale complex is a topological structure defined on scalar functions which extracts critical points of the function and the links between them. By encoding a hierarchy between critical points, less important critical points can be deleted in order to simplify the structure. Our goal is to reconstruct a new shape, which corresponds to the simplified structure while approximating the initial data and preserving the most salient features. We first developed a method for interpolating monotone increasing 2D scalar data with a monotone piecewise cubic C^1-continuous surface. Monotonicity is a sufficient condition for a function to be free of critical points inside its domain. We overcome the restrictive standard axial monotonicity for tensor-product surfaces and introduce sufficient conditions and two algorithms for a more relaxed monotonicity constraint [38], see a piecewise monotonic shape in Figure 6(left). Then, some preliminary results on shape reconstruction from Morse-Smale complexes have been presented as a Posterand at a national conference [35].

In collaboration with Hans Hagen and Anne Berres from University of Kaiserslautern, we investigated conditions under which shape deformations preserve surface curvatures. The work has been published as a chapiter in a scientific book [39].

6.3. Models for Motion Synthesis

- **Scientist in charge**: François Faure.
- **Other permanent researchers**: Marie-Paule Cani, Damien Rohmer, Rémi Ronfard.

Animating objects in real-time is mandatory to enable user interaction during motion design. Physically-based models, an excellent paradigm for generating motions that a human user would expect, tend to lack efficiency for complex shapes due to their use of low-level geometry (such as fine meshes). Our goal is therefore two-folds: first, develop efficient physically-based models and collision processing methods for arbitrary passive objects, by decoupling deformations from the possibly complex, geometric representation; second, study the combination of animation models with geometric responsive shapes, enabling the animation of complex constrained shapes in real-time. The last goal is to start developing coarse to fine animation models for virtual creatures, towards easier authoring of character animation for our work on narrative design.

6.3.1. Real-time physically-based models

Participants: Armelle Bauer, Ali Hamadi Dicko, François Faure, Matthieu Nesme.

Figure 7. Illustration from [14](Left) showing our velocity based adaptive simulation, and from [21](Right) on interactive visualization of muscle activity.

Following the success of frame-based elastic models (Siggraph 2011), a real-time animation framework provided in SOFA and currently used in many of our applications with external partners, we further improved this year the efficiency of this approach: we developed an adaptive version of frame-based elastic models, where frames get seamlessly attached to other ones during deformations when appropriate, in order to reduce computations [14], [33].

Frame-based models were successfully used to model limb movements in anatomical modeling [21]. The efficiency of this mehod enables us to advance towards the concept of a *Living book of anatomy*, where users move their own body and observe it through a tablet to get some visual illustration of anatomy in motion (see Figure 7).

6.3.2. Specific models for virtual creatures

Participants: Marie-Paule Cani, Michael Gleicher.

In collaboration with Loic Barthe and Rodolphe Vaillant from IRIT (U. Toulouse), Brian Wyvill (U. Victoria) and Gael Guennebaud (Manao, Inria), we developed a new automatic method for character skinning: Based on the approximation of character limbs with Hermite RBF implicit volumes, we adjust the mesh vertices representing the skin by projecting them back, at each animation step, to their iso-surface of interest. Since the vertices start from their previous position at the last animation step, there is no need of specifying skinning weights and using another skinning method as pre-computation, as in our previous implicit skinning method. Our solution avoids the well known blending artifacts of linear blend skinning and of dual quaternions, accommodates extreme blending angles and captures elastic effect in skin deformation [16].

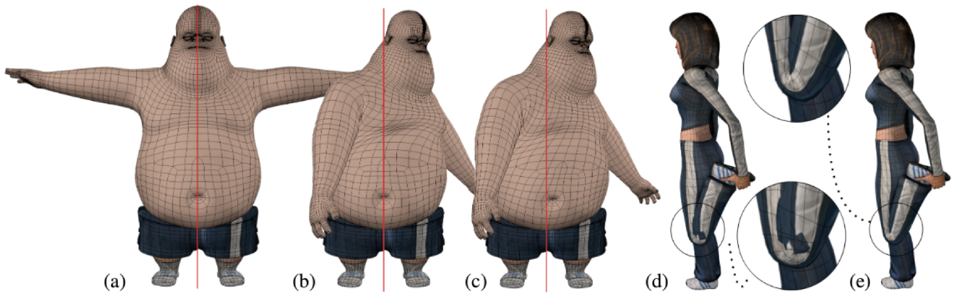


Figure 8. Illustration from [16] showing our results on dynamic skin deformations.

This year, we also studied the way character eyes and gazes are to be animated. This extensive study resulted into a state if the art report published at the Eurographics conference [32]

6.4. Knowledge-based Models for Narrative Design

- **Scientist in charge**: Rémi Ronfard.
- **Other permanent researchers**: Marie-Paule Cani, François Faure, Jean-Claude Léon, Olivier Palombi.

Our long term goal is to develop high-level models helping users to express and convey their own narrative content (from fiction stories to more practical educational or demonstrative scenarios). Before being able to specify the narration, a first step is to define models able to express some a priori knowledge on the background scene and on the object(s) or character(s) of interest. Our first goal is to develop 3D ontologies able to express such knowledge. The second goal is to define a representation for narration, to be used in future storyboarding frameworks and virtual direction tools. Our last goal is to develop high-level models for virtual cinematography such as rule-based cameras able to automatically follow the ongoing action and semi-automatic editing tools enabling to easily convey the narration via a movie.

6.4.1. Knowledge representation through 3D ontologies

Participants: Armelle Bauer, Jean-Claude Léon, Olivier Palombi.

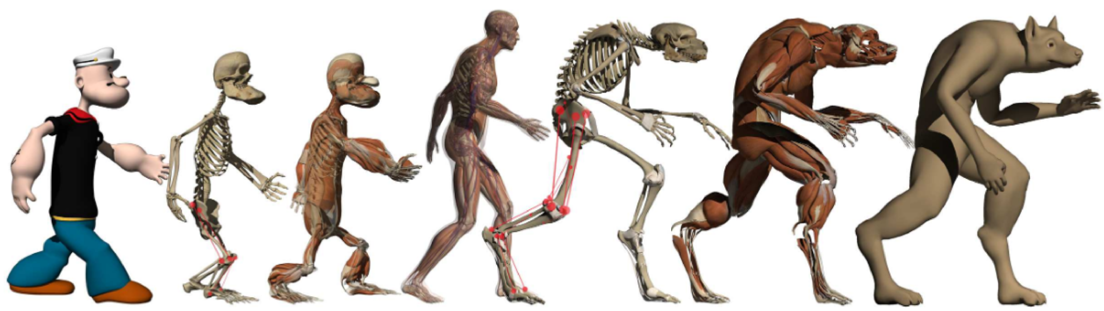


Figure 9. Illustration from [22] showing our results on ontology based anatomy generation.

We chose to develop 3D ontologies for being able to express combined knowledge on geometry, motion and function for assemblies or hierarchies of 3D objects. This is done in collaboration with a specialized group from the LIG laboratory in Grenoble. We decided to first focus these ontologies developments on two topics on which group members have a strong expertise: the anatomical domain (an interesting application test-bed for educational scenarios) and the industrial prototyping domain (where assembly scenarios can be defined).

We developed an anatomical knowledge database called My Corporis Fabrica (MyCF). We first linked functional entities defined in MyCF to the involved anatomical structures, using the musculoskeletal system as a test-bed. Based on this new formal description of the functional anatomy of limbs, we presented a novel pipeline for the construction of biomechanical simulations by combining generic anatomical knowledge with specific data which can handle complex reasoning and querying in MyCF. This resulted into a publication in the Journal of Biomedical Semantics [11]. We also used MyCF within our previous framework of anatomical transfert to set up an assistant tool for modeling and simulating anatomical structure such as bones, muscles, viscera and fat tissues easily while ensuring a correct anatomical consistency [22].

Secondly, in analysing the similarities and differences between existing ontology based describtion of products and virtual humans, we developed a common framework for combining 3D models and functional describtion to both models [15], [34].

6.4.2. Virtual direction tools

Participants: Adela Barbulescu, Rémi Ronfard.

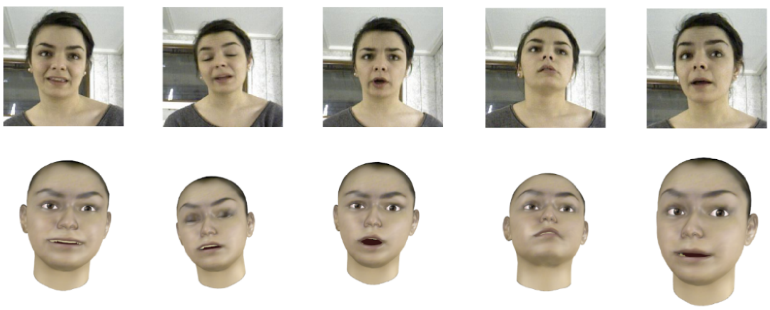


Figure 10. Illustration from [20] showing our results on expressive virtual actors.

We are developing a new approach to transfer speech signals and 3D facial expressions to virtual actors of a different identity. The converted sequences should be perceived as belonging to the target actors. This is the goal of Adela Barbulescu's thesis, co-advised by Gérard Bailly from GIPSA-lab. Our work started with conversion of speaking styles through speech signals only. This year, we started extending this approach to visual prosody and advanced on communicating social attitudes through head gestures [20].

6.4.3. Virtual cinematography

Participants: Quentin Galvane, Vineet Ghandi, Christophe Lino, Rémi Ronfard.

Our goal is to model automatic cameras for covering 3D scenes, as well as to develop semi-automatic film editing techniques to help conveying narration. This work was first conducted on video data, enabling us to test our ideas without the need for complex 3D movies: we designed an automatic method for the identification of actors in a video, and are using it for the automatic re-framing and editing high-resolution videos shots of theater rehearsals [25].

Figure 11. Illustration from [23] showing our results on narrative-driven camera control of computer games.

In parallel, we started extending this methodology to 3D animation, in collaboration with the Mimetic group in Rennes and with Geneva University: this year, we proposed a new method for replaying first person video games with automatic camera control based on the narration [23]. We also advanced towards semi-automatic film editing: A paper was just accepted to AAI 2015. To stress the difficulty of validating film editing methods, we devoted a specific work to validation methodologies [27].

We also addressed other issues related to cinematography and narratives: We designed a pre-visualization system for 3D cinematography to be used in the Action3DS project [30]: the method makes use of 3D modeling to show what the spectators watching a 3D movie are going to see, in order to ease 3D camera control by the film director. Lastly, we worked on computer generation of narrative discourses with the university of Geneva [31].

This year, Remi Ronfard and Vineet Gandhi wrote a patent application "Dispositif de génération de rushes cinématographiques par traitement vidéo", demande de brevet français no. 1460957, déposée le 13 novembre 2014.

6.5. Creating and Interacting with Virtual Prototypes

- **Scientist in charge**: Jean-Claude Léon.
- **Other permanent researchers**: Marie-Paule Cani, Olivier Palombi, Damien Rohmer, Rémi Ronfard.

The challenge is to develop more effective ways to put the user in the loop during content authoring. We generally rely on sketching techniques for quickly drafting new content, and on sculpting methods (in the sense of gesture-driven, continuous distortion) for further 3D content refinement and editing. The objective is to extend these expressive modeling techniques to general content, from complex shapes and assemblies to animated content. As a complement, we are exploring the use of various 2D or 3D input devices to ease interactive 3D content creation.

6.5.1. Sketch-based modeling and editing of 3D shapes

Participants: Marie-Paule Cani, Arnaud Emilien, Even Entem, Stefanie Hahmann, Rémi Ronfard.

While a lot of work has been done on sketch-based modeling of solid shapes, only a few methods do tackle surface models. Terrain surfaces are particular challenging: their fractal-like distribution of details makes them easy to identify, but these cannot be fully drawn by a user. In our work, users only need to draw the main silhouettes they would like to see from a first person viewpoint (enabling, for instance, an art director to set the background scene behind his actors). We generate a plausible, complex terrain that matches the sketch by deforming an existing terrain model. This is done by analyzing the complex silhouettes with cups and T-junctions in the input sketch and matching them with perceptually close features of the input terrain. The rest

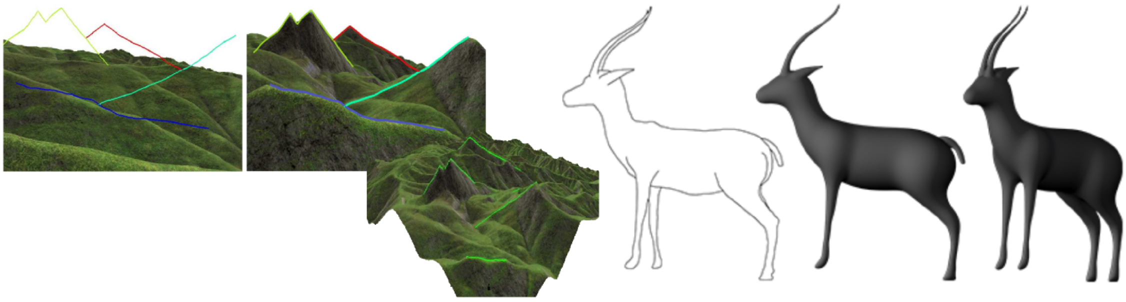


Figure 12. Illustration from [29](Left) showing our results on terrain editing, and from [6](Right) on 3D animal modeling from sketch.

of the terrain is seamlessly deformed while keeping its visual complexity and style. This work was presented at Graphics Interface 2014 [29] and extended to enable the combination of silhouettes from multiple viewpoints in the Computer and Graphics journal [12].

In collaboration with UBC, we introduced the vector graphics complex (VGC), a simple data structure that supports non-manifold topological modeling for 2D vector graphics illustrations. The representation faithfully captures the intended semantics of a wide variety of illustrations, and is a proper superset of scalable vector graphics and planar map representations. VGC nearly separates the geometry of vector graphics objects from their topology, making it easy to deform objects in interesting and intuitive ways, a premise for enabling their animation. This work was published at SIGGRAPH 2014 [4]. We also developed a method for generating 3D animals from a single sketch. This method takes a complex sketch with cups and T-junctions as input (see Figure 12 , and makes use of symmetry hypotheses to analyze it into regions corresponding to the main body and to front and back limbs. The different regions are then automatically reconstructed and blended together using on our implicit modeling methodologies (SCALIS surfaces).

Lastly, we designed a sketch-based interface for authoring illustrative animations. The method makes use of hierarchical motion brushes, a new concept for specifying complex hierarchical motion with a few strokes [28].

6.5.2. Sketching and Sculpting Motion

Participants: Marie-Paule Cani, Arnaud Emilien, Kevin Jordao.

We extended sculpting methods, which had been restricted so far to homogeneous geometric models of a single dimension, to the handling of complex structured shapes and to the interactive sculpting of animated environments.

We developed the first method enabling to sculpt animated content in extending our previous elastic mutable model approach. Relying on the crowd patches representation for modeling animated crowds, we extended component mutations to space-time content, enabling a user to stretch, bend or assemble populated streets while ensuring that individual character trajectories remain continuous through space and time, as well as plausible. This work, developed within Kevin Jordao's thesis co-advised with Julien Pettré from the Mimetic project-team in Rennes, was published at Eurographics 2014 [7].

We developed an interactive system for designing complex waterfall scenes: vector elements created by the user (contacts, freefalls and pools) are used to control the procedural creation of complex waterfalls and rivers that match the user intend while ensuring coherent flows and good embedding within the terrain [18].

6.5.3. Interaction devices and gestural patterns

Participants: Marie-Paule Cani, Rémi Brouet.

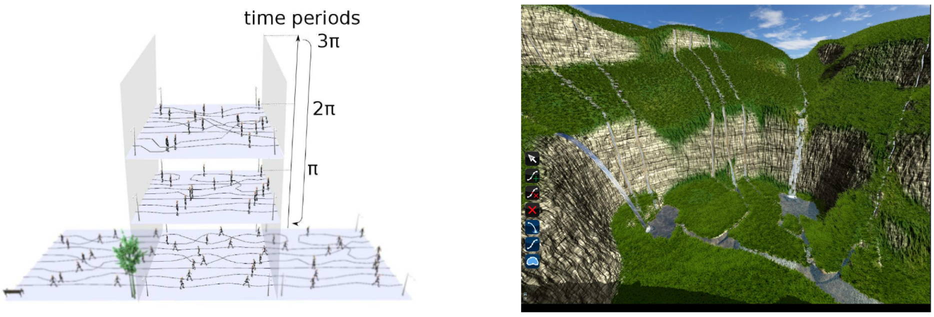


Figure 13. Illustration from [7](Left) on our crowd sculpting interface, and from [18](Right) on interactive design of complex waterfall scenes.

Our work on gestural interaction patterns for 3D design has been developed in collaboration with the HCI team from LIG laboratory towards the exploration of 2D multi-touch tables for the placement and deformation of 3D models.

We are also exploring the use of multi-touch tables for the interactive design and editing of 3D scenes. This is the topic of Rémi Brouet's PhD thesis, co-advised with Renaud Blanch from the human-computer interaction group of LIG laboratory. The main challenge here is to find out how to use 2D interaction media for editing 3D content, hence how to intuitively control the third dimension (depth, non-planar rotations, 3D deformations, etc). Our work on this topic started with a preliminary user study enabling us to analyze all possible hand interaction gestures on table-tops, and to explore the ways users would intuitively try to manipulate 3D environments, either for changing the camera position or for moving objects around. We extracted a general interaction pattern from this study. Our implementation enables both seamless navigation and docking in 3D scenes, without the need for any menu or button to change mode. We are currently extending this work to object editing scenarios, where shapes can be bent or twisted in 3D using 2D interaction and automatic mode selection only.

IN-SITU Project-Team

6. New Results

6.1. Highlights of the Year

Wendy Mackay received the *ACM SIGCHI Lifetime Service Award*.

BEST PAPERS AWARDS :
[22] **CHI '14**. C. LIU, O. CHAPUIS, M. BEAUDOUIN-LAFON, É. LECOLINET, W. E. MACKAY.

6.2. Interaction Techniques

Participants: Caroline Appert, Michel Beaudouin-Lafon, Anastasia Bezerianos, David Bonnet, Olivier Chapuis [correspondant], Cédric Fleury, Stéphane Huot, Can Liu, Justin Mathew, Wendy Mackay, Halla Olafsdottir, Theophanis Tsandilas, Oleksandr Zinenko.

InSitu explores interaction and visualization techniques in a variety of contexts, including individual interaction techniques on different display surfaces that range from mobile devices to very large wall-sized displays, including standard desktop systems and tabletops.

This year, we investigated multi-touch gestures on tabletop [26], we considered the combination of Tilt and Touch on smartphone [29], we proposed novel bi-manual interaction techniques for tablets [18], we introduced a novel focus+context technique to facilitate route following [14], we introduced the *GlideCursor* to facilitate pointing on large display [15], we compared physical navigation in front of a wall-size display with virtual navigation on the desktop [22], we studied users' behavior in immersive Virtual Environments [12], we built a tool to ease the extraction and the expression of parallelism in programs [30] and we investigated the effect of contours on star glyphs [13].

In addition to providing knowledge for designers and practitioners, this set of remarkable results advances our overall knowledge regarding basic interactive phenomena, and allows to better understand how user practices will change.

Multitouch on Tabletop – We systematically studied how users adapt their grasp when asked to translate and rotate virtual objects on a multitouch tabletop [26]. We have shown that users choose a grip orientation that is influenced by three factors: (1) a preferred orientation defined by the start object position, (2) a preferred orientation defined by the target object position, and (3) the anticipated object rotation. We have examined these results in the light of the most recent models of planning for manipulating physical objects and explored how these results can inform the design of tabletop applications.

Tilt & Touch – We studied the combination of tilt and touch when interacting with mobile devices [29]. We conducted an experiment to explore the effectiveness of *TilTouch* gestures for both one-handed and two-handed use. Our results indicate the best combinations of *TilTouch* gestures in terms of performance, motor coordination, and user preferences.

SPad – We created *SPad* [18], a new bimanual interaction technique designed to improve productivity on multi-touch tablets: the user activates quasimodes with the thumb of the non-dominant hand while holding the device with that hand and interacts with the content with the dominant hand (figure 3). We conducted an iterative design process and created a tablet application that demonstrates how SPad enables faster, more direct and more powerful interaction without increasing complexity.

Figure 3. Left: SPad control accessible with the thumb. Three commands are accessible with a tap of the thumb, 4 menus are accessible with swipes. Right: SPad in use to paste and move objects that have just been copied.

RouteLenses – Millions of people go to the Web to search for geographical itineraries. Inspecting those map itineraries remains tedious because they seldom fit on screen, requiring much panning & zooming to see details. Focus+context techniques address this problem by displaying routes at a scale that allows them to fully fit on screen: users see the entire route at once, and perform magnified steering using a lens to navigate along the path, revealing additional detail. We created *RouteLenses* [14], a type of lenses that automatically adjusts their position based on the geometry of the path that users steer through (figure 4 . RouteLenses make it easier for users to follow a route, yet do not constrain movements too strictly, leaving them free to move the lens away from the path to explore its surroundings.

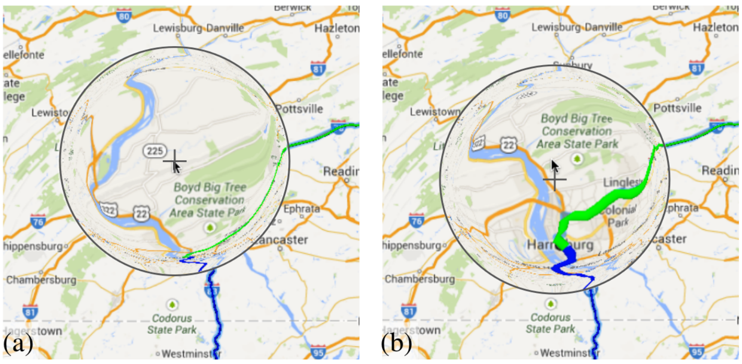


Figure 4. Following an itinerary. (a) Conventional lens: the user overshoots at a right turn in Harrisburg; losing the route that falls in the distorted region. (b) RouteLens: the route's attraction compensates the overshoot; the lens remains closer to the route, which remains in focus.

GlideCursor – Pointing on large displays with an indirect, relative pointing device such as a touchpad often requires clutching. We designed and evaluated *GlideCursor* [15], which lets the cursor continues to move during clutching gestures. The effect is that of controlling the cursor as a detached object that can be pushed, with inertia and friction similar to a puck being pushed on a table. We analyzed gliding from a practical and a theoretical perspective and conducted two studies. The first controlled experiment established that gliding reduces clutching and can improve pointing performance for large distances. We introduced a measure called *cursor efficiency* to capture the effects of gliding on clutching. The second experiment demonstrated that participants use gliding even when an efficient acceleration function lets them perform the task without it, without degrading performance.

Wall vs. Desktop – The advent of ultra-high resolution wall-size displays and their use for complex tasks require a more systematic analysis and deeper understanding of their advantages and drawbacks compared with desktop monitors. While previous work has mostly addressed search, visualization and sense-making tasks, we have designed and evaluated an abstract classification task that involves explicit data manipulation [22]. Based on our observations of real uses of a wall display (figure 5 -left), this task represents a large category of applications. We conducted a controlled experiment that uses this task to compare physical navigation in front of a wall-size display (figure 5 -right) with virtual navigation using pan-and-zoom on the desktop. Our main finding is a robust interaction effect between display type and task difficulty: while the desktop can be faster than the wall for simple tasks, the wall gains a sizable advantage as the task becomes more difficult.

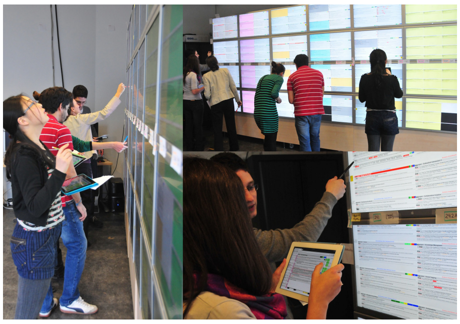


Figure 5. Left: Fine tuning the CHI 2013 conference schedule on the WILD display. Center: Classification task inspired by the CHI scheduling task conducted on the wall-sized display. Subjects must move misclassified red disks into containers (represented by the individual screens) of the same class. The class is represented by a very small letter at the center of each disk, forcing subjects to physically move in front of the display.

Immersive VE – The feeling of presence is essential for efficient interaction within Virtual Environments (VEs). When a user is fully immersed within a VE through a large immersive display system, her feeling of presence can be altered because of disturbing interactions with her physical environment. This alteration can be avoided by taking into account the physical features of the user as well as those of the system hardware. Moreover, the 3D abstract representation of these physical features can also be useful for collaboration between distant users. In [12] we presented how we use the Immersive Interactive Virtual Cabin (IIVC) model to obtain this virtual representation of the user's physical environment and we illustrated how this representation can be used in a collaborative navigation task in a VE. We also presented how we can add 3D representations of 2D interaction tools in order to cope with asymmetrical collaborative configurations, providing 3D cues for a user to understand the actions of others even if he/she is not fully immersed in the shared VE.

Clint – We created *Clint*, a direct manipulation tool to ease the extraction and the expression of parallelism in existing programs [30]. Clint is built on top of state-of-the-art compilation tools (polyhedral representation of programs) in order to give a visual representation of the code, perform automatic data dependence analysis and to ensure the correctness of code transformations (figure 6). It can be used to rework and improve automatically generated optimizations and to make manual program transformation faster, safer and more efficient.

Start Glyphs – We conducted three studies using crowd-sourcing on Amazon mechanical Turk, to determine the effect of using contours on data glyphs such as star glyphs [13]. Our results indicate that glyphs without

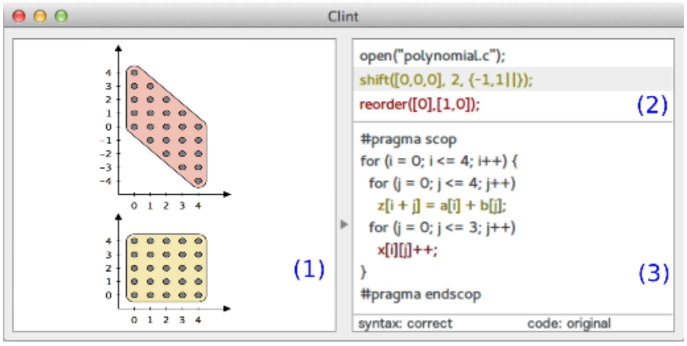


Figure 6. Clint interface includes: (1) interactive visualization, (2) editable history view, and (3) source code editor.

contours lead viewers to naturally make judgements that are data-driven. Whereas adding contours encourages shape similarity, e.g. perceiving rotated variations of glyphs as similar (even though they are not similar in data space).

6.3. Research Methods

Participants: Michel Beaudouin-Lafon, Anastasia Bezerianos, Jérémie Garcia, Stéphane Huot, Ilaria Liccardi, Wendy Mackay [correspondant], Justin Mathew.

Conducting empirical research is a fundamental part of InSitu's research activities, including observation of users in field and laboratory settings to discover problems faced by users, controlled laboratory experiments to evaluate the effectiveness of the technologies we develop, longitudinal field studies to determine how our technologies work in the real world, and participatory design, to explore design possibilities with users throughout the design process.

Computer-aided Composition – We designed *Polyphony* [20], a novel interface for systematically studying all phases of computer-aided composition, and then used it to observe expert creative behavior. *Polyphony* is a unified user interface that integrates interactive paper and electronic user interfaces for composing music. We asked 12 composers to use it (figure 7 -left) to compose an electronic accompaniment to a 20-second instrumental composition by Anton Webern. The resulting dozen comparable snapshots of the composition process reveal how composers both adapt and appropriate tools in their own way. In collaboration with IRCAM, we also conducted a longitudinal study where we closely collaborated with composer Philippe Leroux [19] in the creation of his piece *Quid sit musicus*. The composer used our interfaces based on interactive paper along with an OpenMusic library to generate compositional material for this work (figure 7 -right).

Multitouch Gestures – We created a design space of simple multitouch gestures that designers of user interfaces can systematically explore to propose more gestures to users [27]. We further considersed a set of 32 gestures for tablet-sized devices, by developing an incremental recognition engine that works with current hardware technology, and empirically testing the usability of those gestures. In our experiment, individual gestures were recognized with an average accuracy of ∼90%, and users successfully achieved some of the transitions between gestures without the use of explicit delimiters. The goal of this work is to assist designers in optimizing the use of the rich multi-touch input channel for the activation of discrete and continuous controls, and to enable fluid transitions between controls, e.g. when selecting text over multiple views, manipulating different degrees of freedom of a graphical object or invoking a command and setting its parameter values in a row.

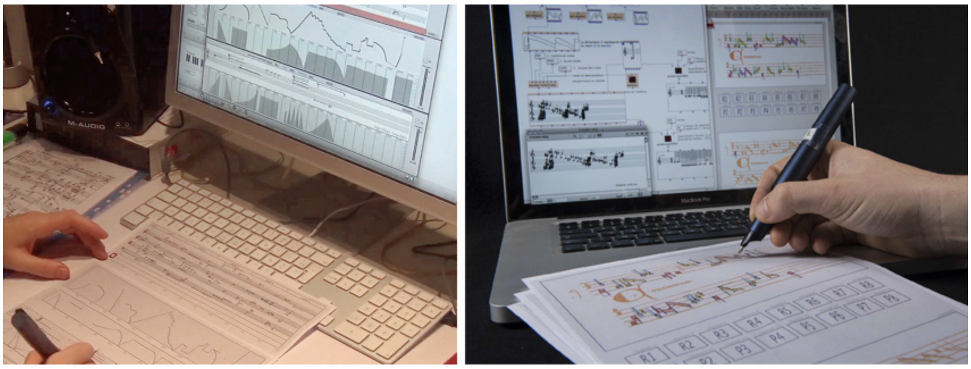


Figure 7. Left: A composer completes the composition task using Polyphony. Right: Set of tools created for Philippe Leroux' piece Quid sit musicus. Photo by H. Raguet, Inria.

Spatial Audio – We investigated the issues of spatialization techniques for object-based audio production and introduced the Spatial Audio Design Spaces framework (SpADS) [25], which describes the spatial manipulation of object-based audio. These design spaces are based on interviews with professional sound engineers and on a morphological analysis of 3D audio objects that clarifies the relationships between recording and rendering techniques that define for 3D speaker configurations. This will allow us to analyze and design novel advanced object-based controllers.

Physical Visualizations – We studied the design process of physical visualizations. An increasing variety of such visualizations are being built, for purposes ranging from art and entertainment to business analytics and scientific research. However, crafting them remains a laborious process and demands expertise in both data visualization and digital fabrication. We analyzed the limitations of current workflows through three real case studies and created *MakerVis*, the first tool that integrates the entire workflow, from data filtering to physical fabrication (figure 8). Design sessions with three end users showed that tools such as *MakerVis* can dramatically lower the barriers behind producing physical visualizations. Observations and interviews also revealed important directions for future research. These include rich support for customization, and extensive software support for materials that accounts for their unique physical properties as well as their limited supply.

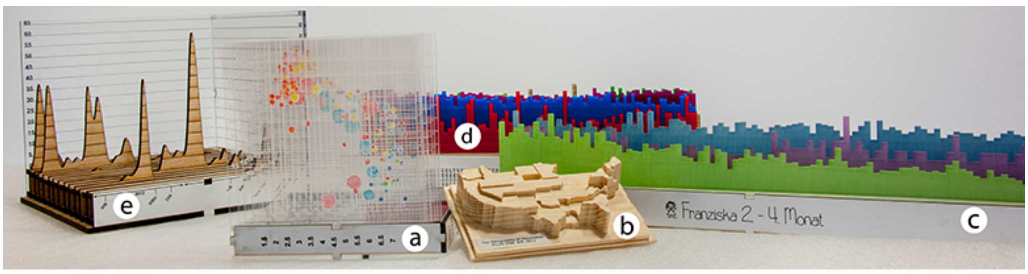


Figure 8. Physical visualizations created with MakerVis: a) a scatterplot created after Hans Rosling's TED talk, b) a prism map showing happiness across the US computed from Twitter sentiments, c),d),e) visualizations created by users during design sessions.

6.4. Engineering of interactive systems

Participants: Caroline Appert, Michel Beaudouin-Lafon [correspondant], Olivier Chapuis, Cédric Fleury, Stéphane Huot, Theophanis Tsandilas, Wendy Mackay.

InSitu has a long tradition of developing software tools and user interface toolkits to facilitate the creation of interactive systems. These tools allow us to better experiment with our ideas and are therefore an integral part of our research methodology. Most of them are freely available and some are used outside InSitu for research or teaching.

Interactive Paper – We created *PaperComposer* [31], a graphical interface builder for creating personal interactive-paper applications for musical creation. We also built an API that facilitates the development of interactive paper components for PaperComposer. The API enables developers to define new paper components that accept additional musical data with their own representation structures and interactions.

3D Telepresence – In the context of 3D telepresence, we studied how to transmit a 3D model of the users to a remote location. In [17] we present a 3D head reconstruction method for low cost 3D telepresence systems that uses only a single consumer level hybrid sensor (color+depth) located in front of the users. Our method fuses the real-time, noisy and incomplete output of a hybrid sensor with a set of static, high-resolution textured models acquired in a calibration phase (figure 9). A complete and fully textured 3D model of the users' head can thus be reconstructed in real-time, accurately preserving the facial expression of the user. The main features of our method are a mesh interpolation and a fusion of a static and a dynamic textures to combine respectively a better resolution and the dynamic features of the face.

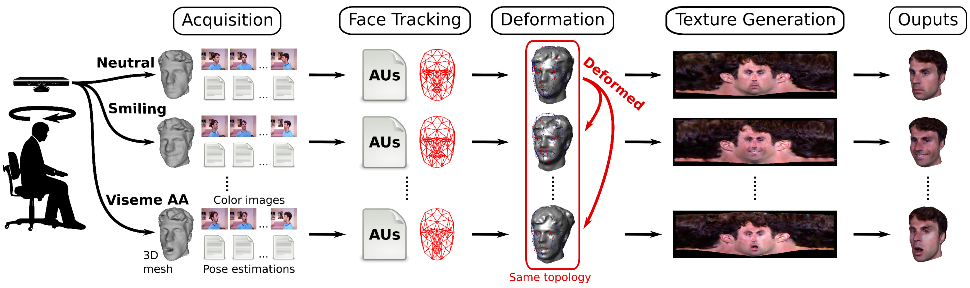


Figure 9. Acquisition step of the 3D face reconstruction: data are processed to create a complete and fully textured 3D head model for each facial expression. This set of head models are then used to improve the real-time reconstruction of the user's head during a 3D telepresence session.

Wall-sized displays – We developed Smarties [16], a system that allows developers to easily add interactive support to their wall-sized display applications by using mobile devices such as tablets. The system includes an original mobile interface that can be customized by the application itself (without programming the mobile device), a communication protocol between the mobile devices and the application running on the wall-sized display, and libraries in different programming languages that implement the protocol and handle synchronization, locking and input conflicts. Synchronization between multiple mobile devices is handled by the libraries, and thus the system supports free collaboration. The mobile devices come with multiple cursor controllers, also associated with keyboards, widgets and clipboards.

INDES Project-Team

6. New Results

6.1. Web programming

Participants: Yoann Couillec, Vincent Prunet, Tamara Rezk, Manuel Serrano [correspondant].

6.1.1. Hop.js

Multitier programming languages unify within a single formalism and a single execution environment the programming of the different tiers of distributed applications. On the Web, this programming paradigm unifies the client tier, the server tier, and, when one is used, the database tier. This homogenization offers several advantages over traditional Web programming that rely on different languages and different environments for the two or three tiers of the Web application: programmers have only one language to learn, maintenance and evolution are simplified by the use of a single formalism, global static analyses are doable as a single semantics is involved, debugging and other runtime tools are more powerful as they access global informations about the execution [17].

The three first multitier platforms for the Web all appeared in 2006: GWT (a.k.a., Google Web Toolkit), Links, and Hop [6], [5]. Each relied on a different programming model and languages. GWT maps the Java programming model on the Web, as it allows, Java/Swing likes programs to be compiled and executed on the Web; Links is functional language with experimental features such as the storing of the whole execution context on the client; Hop is based on the Scheme programming language. These three pioneers have open the path for the other multitier languages such as, Ocsigen for Ocaml, UrWeb, js-scala, etc.

In spite of their interesting properties, multitier languages have not become that popular on the Web. Today, only GWT is widely used in industrial applications but arguably GWT is not a fully multitier language as developing applications with GWT requires explicit JavaScript and HTML programming. This lack of popularity of other systems is likely due to their core based languages than to the programming model itself.

JavaScript is the *defacto* standard on the Web. Since the mid 90's, it is the language of the client-side programming and more recently, with systems like nodejs, it is also a viable solution for the server-side programming. As we are convinced by the virtues of multitier programming we have started a new project consisting of enabling multitier programming JavaScript. We have created a new language called HopScript, which is a minimalist extension of JavaScript for multitier programming, and we have implemented a brand new runtime environment called Hop.js. This environment contains a builtin Web server, on-the-fly HopScript compilers, and many runtime libraries.

HopScript is a super set of JavaScript, *i.e.,* all JavaScript programs are legal HopScript programs. Hop.js is a compliant JavaScript execution environment as it succeeds at 99% of the Ecma 262 tests suite. The Hop.js environment also aims at Node.js compatibility. In its current version it supports about 70% of the Node.js runtime environment. In particular, it fully supports the Node.js modules, which lets Hop programs reuse existing Node.js modules as is.

A prototype version of Hop.js is currently used by several academic and SME R&D teams to jointly develop an assistive robotic platform and a set of distributed applications.

We plan to release the first public Hop.js version by the end of the first semester of 2015, as we plan to start describing in forthcoming papers.

6.1.2. Multitier Debugging

Debugging Web applications is difficult because of their distributed nature and because the server-side and the client-side of the application are generally treated separately. The multitier approach, which reunifies the two ends of the application inside a unique execution environment, helps the debugging process because it lets the debugger access more runtime informations.

Based on our previous work on the Hop multitier debugger [17], we have built a multitier debugger for Hop.js, our multitier extension of JavaScript. Its advantage over most debuggers for the Web is that it reports the full stack trace containing all the server-side and client-side frames that have conducted to an error. Errors are reported on their actual position on the source code, wherever they occur on the server or on the client. This paper presents this debugger and sketches its implementation. This work is described in a yet unpublished paper, which will appear in 2015.

6.1.3. Datasource

We extended the HOP.JS language with an embedded language, inspired by PLINQ and ORC, called DATASOURCE. It allows programmers request multiple data sources with queries written in a unique language. We used a plinq-like language to express queries and an orc-like language to orchestrate them. Our query language and the orchastration languages can be used simultaneously or separately. We implemented bindings between DATASOURCE and some representative types of data sets such as SPARQL endpoints, relationnal databases, WEB services, and WEB pages. We are extending HOP.JS by supporting EcmaScript 6 array comprehensions in order to write a unique query over multiple data sources in a unified formalism. The query is then compiled into database specific queries. We linked all the bindings made for HOP with HOP.JS. We implemented another binding for a document oriented data base, MONGODB.

6.2. Distributed programming

Participant: Bernard Serpette [correspondant].

6.2.1. Logical behavioural semantics of Esterel

We have formalised, with the Coq system, the logical behavioural semantics of Esterel as described in Gérard Berry's book. In order to define the properties of reactivity and determinism, we have defined a new semantics using contexts with a proven correspondence between the two semantics.

The specification and the proofs of the correspondence take 3500 lines of Coq.

6.2.2. Abstract distributed machine

We have experimented an abstract machine composed of distributed nodes. Each node has exactly two named links to other nodes and an instruction able to modify one link of a reachable node. This instruction is executed when a token is received, once the instruction is achieved the token is transmitted to another reachable node.

This abstract machine is turing complete. The λ-calculus and the π-calculus can be compiled to the instruction set of this machine.

The execution of one individual node may involve paths of arbitrary length, for example, when compiling the λ-calculus or the π-calculus, the path length for accessing a variable is proportional to it's de Bruijn index and therefore is not bounded. Given a machine with instructions of unbounded paths, we can build an equivalent machine where all the paths are bounded by two: a node is only able to access it's own links and the links of its neighbour. Moreover, this transformation uses only 6 different instructions.

6.3. Security and Privacy

Participants: Ilaria Castellani, José Fragoso Santos, Nataliia Bielova, Tamara Rezk [correspondant].

6.3.1. Security of Dynamically Evolving Systems of Communicating Processes

We have started to address security issues in the context of dynamically evolving systems of communicating processes, which are able to adapt themselves in reaction to particular events (for instance, security attacks or changes in security policies). We present initial results on a simple model of processes communicating via structured interactions (sessions), in which self-adaptation and security concerns are jointly addressed. In this model, security violations occur when processes attempt to read or write messages of inappropriate security level within a structured interaction. Such violations trigger adaptation mechanisms that prevent the violations to occur and/or to propagate their effect in the choreography. Our model is equipped with local and global mechanisms for reacting to security violations; type soundness results ensure that the global protocols are still correctly executed while the system adapts itself to preserve its security.

6.3.2. Browser Randomisation against Fingerprinting: a Quantitative Information Flow Approach

Web tracking companies use device fingerprinting to distinguish the users of the websites by checking the numerous properties of their machines and web browsers. One way to protect the users' privacy is to make them switch between different machine and browser configurations. We propose a formalisation of this privacy enforcement mechanism.

We use information-theoretic channels to model the knowledge of the tracker and the fingerprinting program, and show how to synthesise a randomisation mechanism that defines the distribution of configurations for each user. This mechanism provides a strong guarantee of *privacy* (the probability of identifying the user is bounded by a given threshold) while maximising *usability* (the user switches to other configurations rarely). To find an optimal solution, we express the enforcement problem of randomisation by a linear program. We investigate and compare several approaches to randomisation and find that more efficient privacy enforcement would often provide lower usability. Finally, we relax the requirement of knowing the fingerprinting program in advance, by proposing a randomisation mechanism that guarantees privacy for an arbitrary program.

This work has been published and presented at the Nordic Conference on Secure IT Systems (NordSec 2014) [12]. The extended version of the paper has been published as a technical report [20].

6.3.3. Crying Wolf? On the Price Discrimination of Online Airline Tickets

Price discrimination refers to the practice of dynamically varying the prices of goods based on a customer's purchasing power and willingness to pay. Motivated by several anecdotal accounts, we report on a three week experiment, conducted in search of price discrimination in airline tickets. Despite presenting the companies with multiple opportunities for discriminating us, and contrary to our expectations, we did not find any evidence for systematic price discrimination. At the same time, we witnessed the highly volatile prices of certain airlines which make it hard to establish cause and effect. Finally, we provided alternative explanations for the observed price differences.

This work has been published and presented at the Workshop on Hot Topics in Privacy Enhancing Technologies (HotPETs 2014) [19].

6.3.4. Stateful Declassification Policies for Event-Driven Programs

We propose a novel mechanism for enforcing information flow policies with support for declassification on event-driven programs. Declassification policies consist of two functions. First, a projection function specifies for each confidential event what information in the event can be declassified directly. This generalizes the traditional security labelling of inputs. Second, a stateful release function specifies the aggregate information about all confidential events seen so far that can be declassified. We provide evidence that such declassification policies are useful in the context of JavaScript web applications. An enforcement mechanism for our policies is presented and its soundness and precision is proven. Finally, we give evidence of practicality by implementing and evaluating the mechanism in a browser. This work has been published at Computer Security Foundations (CSF'14) [18].

6.3.5. An Information Flow Monitor for a Core of DOM

We propose and prove sound a novel, purely dynamic, flow sensitive monitor for securing information flow in an imperative language extended with DOM-like tree operations, that we call Core DOM. In Core DOM, as in the DOM API, tree nodes are treated as first-class values. We take advantage of this feature in order to implement an information flow control mechanism that is finer-grained than previous approaches in the literature. Furthermore, we extend Core DOM with additional constructs to model the behavior of live collections in the DOM Core Level 1 API. We show that this kind of construct effectively augments the observational power of an attacker and we modify the proposed monitor so as to tackle newly introduced forms of information leaks. This work has been published at the 9th International Symposium on Trustworthy Global Computing (TGC) [11].

6.3.6. An Information Flow Monitor-Inlining Compiler for Securing a Core of JavaScript

Web application designers and users alike are interested in isolation properties for trusted JavaScript code in order to prevent confidential resources from being leaked to untrusted parties. Noninterference provides the mathematical foundation for reasoning precisely about the information flows that take place during the execution of a program. Due to the dynamicity of the language, research on mechanisms for enforcing noninterference in JavaScript has mostly focused on dynamic approaches. We present the first information flow monitor inlining compiler for a realistic core of JavaScript. We prove that the proposed compiler enforces termination-insensitive noninterference and we provide an implementation that illustrates its applicability.

This work has been published at the 29th IFIP International Information Security and Privacy Conference (IFIP SEC) [14].

6.3.7. From Static to Hybrid Typing Secure Information Flow in a Core of JavaScript

We propose a novel type system for securing information flow in a core of JavaScript. This core takes into account the defining features of the language, such as prototypical inheritance, extensible objects, and constructs that check the existence of object properties. We design a hybrid version of the proposed type system. This version infers a set of assertions under which a program can be securely accepted and instruments it so as to dynamically check whether these assertions hold. By deferring rejection to runtime, the hybrid version can typecheck secure programs that purely static type systems cannot accept.

INFINE Team

6. New Results

6.1. Highlights of the Year

- We proved a conjecture made in 2011 about the feasibility of non-trivial community detection just above a threshold below which it was known that only trivial detection could be done, see [13]. This was published in ACM STOC'14 and well-received, as the proof required the invention of new techniques to control the spectral properties of random matrices.
- The official opening of IoT-LAB of all sites through the "Workshop Internet Of Things/Equipex FIT IoT-LAB" held in Grenoble (on 6 and 6 november 2014), has been a major event for our team: it concludes several years of preparation of the IoT-LAB site located in Rocquencourt, currently managed by C. Adjih, E. Baccelli and I. Amdouni, which was itself opened the same month https://www.iot-lab.info/opening-of-the-paris-rocquencourt-site/.

6.2. Panorama

All the INFINE research activities encompass both theoretical or protocol designing research (to seek for conceptual advances or optimizations) and applied research (to validate and/or experiment the proposed concepts against real networking scenarios). The target applications range from Internet-based applications to mobile wireless networks. Taking an information- and user-centric perspective, we envision networks as means to convey relevant information to users, while adapting to customary practices (in terms of context, interests, or content demands) of such users. INFINE is thus organized along three main axes, namely Online Social Networks, Resource and Traffic Management, and Spontaneous Wireless Networks.

6.3. Online Social Networks (OSN)

Community detection; bandit algorithms; privacy preservation; reward mechanisms

6.3.1. Community detection

Participants: Laurent Massoulié, Marc Lelarge, Jiaming Xu.

We have progressed in the design of spectral methods for community detection and in the corresponding analysis (see above and references [3], [13], [22]).

6.3.2. Bandit algorithms for active learning of content type at low spam cost

Participants: Laurent Massoulié, Mesrob Ohanessian, Alexandre Proutière.

We developed a framework in which to cast the problem, and the so-called “greedy Bayes” algorithm to determine which user to expose to a given content. We proved corresponding optimality properties, and observed that “greedy Bayes” beats the so-called Thompson sampling approach, that is the state-of the-art method in bandit problems. Work currently under submission.

6.4. Resource and Traffic Management

Traffic offloading; infrastructure deployment; opportunistic routing; traffic modeling; intermittently connected networks.

6.4.1. From Routing to Network Deployment for Data Offloading in Metropolitan Areas

Participants: Eduardo Mucceli, Aline Carneiro Viana.

Smartphone sales are booming, nearly half billion were sold in 2011; more smartphones, more mobile data traffic, and Currently, 3G cellular networks in metropolitan areas are struggling to attend the recent boost up of mobile data consumption. Carefully deploying WiFi hotspots allow to maximize WiFi offloading and can both be cheaper than upgrade the cellular network structure and concede substantial improvement in the network capacity. In this context, in this work, we first propose a new way to map into a graph the *people behavior* (i.e., mobility context) in an urban scenario. Our proposed behavior-to-graph solution is simple, take into consideration the restrictions imposed by transportation modes to traffic demand, the space-time interaction between people and urban locations, and finally, is powerful to be used as input to any popular area identification problem (key points for an efficient network planning). Secondly, we propose a metric to identify locations more capable of providing coverage for people and consequently, more suitable for receiving hotspots. Deploying a small percentage of hotspots ranked by the herein proposed metric provides high percentages of coverage time for people moving around in the city. Using a real-life metropolitan trace, we show our routine-based strategy guarantees higher offload ratio than current approaches in the literature while using a realistic traffic model. Different parts of this work has been published in the international conferences IEEE SECON 2014 [14], IEEE WCNC 2014 [18] and IEEE WMNC 2014 [17], and in the international Student workshop IEEE Infocom [15]. An extended version of this paper is under submission in a trasaction. This version includes new characterization results of the used trace and new analysis of space-traffic correlation.

6.4.2. Mobile Data Traffic Modeling: Revealing Hidden Facets

Participants: Eduardo Mucceli, Aline Carneiro Viana, Kolar Purushothama Naveen, Carlos Sarraute.

Smartphone devices provide today the best means of gathering users information about content consumption behavior on a large scale. In this context, the literature is rich in work studying and modeling users mobility, but little is publicly known about users content consumption patterns. The *understanding of users' mobile data traffic demands* is of fundamental importance when looking for solutions to manage the recent boost up of mobile data usage [14] and to improve the quality of communication service provided. Hence, the definition of a *usage pattern* can allow telecommunication operators to better forseen future demanded traffic and consequently, to better (1) deploy data offloading hotspots or (2) timely plan network resources allocation and then, set subscription plans.

Using a large-scale dataset collected from a major 3G network in a big metropolitan area, in this work, we present the first detailed measurement-driven modeling of mobile data traffic usage of smartphone subscribers. Our main outcome is a synthetic measurement-based mobile data traffic generator, capable of imitating traffic-related activity patterns of different categories of subscribers and time periods of a routinary normal day in their lives. For this, we first characterize individual subscribers routinary behaviour, followed by the detailed investigation of subscribers' usage pattern (i.e., "when" and "how much" traffic is generated). Broadly, our observations bring important insights into network resource usage. We then classify the subscribers into six distinct profiles according to their usage pattern and model these profiles according to two different journey periods: peak and non-peak hours. We show that the synthetic trace generated by our data traffic model consistently imitates different subscriber profiles in two journey periods, when compared to the original dataset. We discuss relevant issues in traffic demands and describe implications in network planning and privacy. This work has been published in the international conference IEEE PERCOM 2014 [16]. An extended version of this paper is under submission in a trasaction and a technical report is available in [26]. This version includes new characterization results of the used trace, including analysis correlating age and gender to traffic demands, as well as new profiling results.

6.4.3. On the Interaction between Content Caching and Routing

Participants: Kolar Purushothama Naveen, Laurent Massoulié, Emmanuel Baccelli, Aline Carneiro Viana, Don Towsley.

Nowadays Internet users are mobile over 60% of their time online, and mobile data traffic is expected to increase by more than 60% annually to reach 15.9 exabytes per month by 2018. This evolution will likely incur durably congested wireless access at the edge despite progress in radio technologies. To alleviate congestion at the Internet edge, one promising approach is to target denser deployments of wireless access points. As a result, mobile users are potentially within radio reach of several access points (AP) from which content may be directly downloaded. In this context, distinct AP's can have very different bandwidth and memory capacities. Such differences raise the following question: When requests can be sent to several such access points, how to optimize performance through both load balancing and content replication?

In this work, we introduce formal optimization models to address this question, where bandwidth availability is represented via a cost function, and content availability is represented either by a cost function or a sharp constraint. For both formulations we propose dynamic caching and request assignment algorithms. Crucially our request assignment scheme is based on a server price signal jointly reflecting content and bandwidth availability. Using mean field approximation and Lyapunov functions techniques, we prove that our algorithms are optimal and stable in a limiting fluid regime with large arrival rates and content chunking. Through simulations we exhibit the efficacy of our request assignment strategy in comparison to the common practices of assigning requests purely based on either bandwidth or content availability. Finally, using the popular LRU (Least Recently Used) strategy instead for cache replacements, we again demonstrate the superior performance of our request assignment strategies. This work in under submission in an international conference.

6.4.4. Data Delivery in Opportunistic and Intermittently Connected Networks

Participants: Ana Cristina Vendramin, Anelise Munaretto, Myriam Delgado, Aline Carneiro Viana, Mauro Fonseca.

The pervasiveness of computing devices and the emergence of new applications and cloud services are factors emphasizing the increasing need for adaptive networking solutions. In most cases, this adaptation requires the design of interdisciplinary approaches as those inspired by nature, social structures, games, and control systems. The approach presented in this work brings together solutions from different, yet complementary domains, i.e., networking, artificial intelligence, and complex networks, and is aimed at addressing the problem of efficient data delivery in intermittently connected networks.

As mobile devices become increasingly powerful in terms of communication capabilities, the appearance of opportunistic and intermittently connected networks referred to as Delay Tolerant Networks (DTNs) is becoming a reality. In such networks, contacts occur opportunistically in corporate environments such as conferences sites, urban areas, or university campuses. Understanding node mobility is of fundamental importance in DTNs when designing new communication protocols that consider opportunistic encounters among nodes. This work proposes the Cultural Greedy Ant (CGrAnt) protocol to solve the problem of data delivery in opportunistic and intermittently connected networks. CGrAnt is a hybrid Swarm Intelligence-based forwarding protocol designed to address the dynamic and complex environment of DTNs. CGrAnt is based on: (1) Cultural Algorithms (CA) and Ant Colony Optimization (ACO) and (2) operational metrics that characterize the opportunistic social connectivity between wireless users. The most promising message forwarders are selected via a greedy transition rule based on local and global information captured from the DTN environment. Using simulations, we first analyze the influence of the ACO operators and CA knowledge on the CGrAnt performance. We then compare the performance of CGrAnt with the PROPHET and Epidemic protocols (two well known related protocols in the literature) under varying networking parameters. The results show that CGrAnt achieves the highest delivery ratio (gains of 99.12% compared with PROPHET and 40.21% compared with Epidemic) and the lowest message replication (63.60% lower than PROPHET and 60.84% lower than Epidemic). This work is under submission to an international journal. Some parts of this work were previously published in the international conference ACM GECCO 2012 and in the Elsevier Computer Networks journal.

6.4.5. Vehicular Network under a Social Perception

Participants: Felipe D. Cunha, Aline Carneiro Viana, Raquel A. F. Mini, Antonio A.f. Loureiro.

Vehicular Mobility is strongly influenced by the speed limits, destinations, traffic conditions, period of the day, and direction of the public roads. At the same time, the driver's behavior produces great influences in vehicular mobility. People tend to go to the same places, at the same day period, through the same trajectories, which le ad them to the appearance of driver's daily routines. These routines lead us to the study of mobility in VANETs under a social perspective and to investigate how effective is to explore social interactions in this kind of network. In this work, we thus characterize and evaluate social properties of a realistic vehicular trace found in literature. Our aim is to study the vehicles' mobility in accordance to social behaviors. Social metrics are computed and the obtained results are compared to random graphs. With our analysis, we could verify the existence of regularity and common interests among the drivers in vehicular networks. This work was published in the international conference IEEE ISCC 2014 [10], in the international Student workshop of IEEE Infocom 2014 [9], and at the internation workshop Internet of Things Communications and Technologies (IoT 2013) held in conjunction with IEEE WiMob 2013.

After having identified routine in vehicles mobility patterns and their correlation with the period of the day, we then leverage the identified social aspects to design a *Socially Inspired Broadcast Data Dissemination* for VANETs. We claim that protocols and applications designed for Vehicular Ad Hoc Networks need to adapt to vehicles routines in order to provide better services. With this issue in mind, we designed a data dissemination solution for these networks that considers the daily road traffic variation of large cities and the relationship among vehicles. The focus of our approach is to select the best vehicles to rebroadcast data messages according to social metrics, in particular, the clustering coefficient and the node degree. Moreover, our solution is designed in such a way that it is completely independent of the perceived road traffic density. Simulation results show that, when compared to related protocols, our proposal provides better delivery guarantees, reduces the network overhead and possesses an acceptable delay. This work was published as a short paper at the international conference ACM MSWiM 2014 [8]

6.4.6. Design and Analysis of an Efficient Friend-to-Friend Content Dissemination System

Participants: Kanchana Thilakarathna, Aline Carneiro Viana, Aruna Seneviratne, Henrik Petander.

In this work, we focus on dissemination of content for delay tolerant applications/services, (i.e. content sharing, advertisement propagation, etc.) where users are geographically clustered into communities. Due to emerging security and privacy concerns, majority of users are becoming more reluctant to interact with strangers and are only willing to share information/content with the users who are previously identified as friends. As a result, despite its promise, opportunistic communications systems have not been widely adopted. In addition, in this environment, opportunistic communication will not be effective due to the lack of known friends within the communication range. We thus propose a novel architecture which combines the advantages of distributed decentralized storage and opportunistic communications. The proposed system addresses the trust and privacy concerns of opportunistic communications systems, and enables the provision of efficient distributed mobile social networking services. We exploit the fact that users will trust their friends, and the friends will help in disseminating content by temporarily storing and forwarding content. This can be done by replicating content on friends' devices who are likely to consume that content and provide the content to other friends when the device has access to low cost networks. The fundamental challenge then is to minimize the number of replicas, to ensure high and timely availability. We provide a formal definition of this content replication problem, and show that it is NP hard. Then, we propose a community based greedy heuristic algorithm with novel dynamic centrality metrics that replicates the content on a minimum number of friends' devices, and maximizes the availability of content. Using both real world and synthetic traces, we validate effectiveness of the proposed scheme. In addition, we demonstrate the practicality of the the proposed system, through an implementation on Android smartphones. This work is under submission in an international transaction. An initial version of this work was published at the international conference ACM MobiHoc 2013, and an extended version is under submission in an international trasaction.

6.4.7. Telling Apart Social and Random Relationships in Dynamic Networks

Participants: Pedro Olmo Vaz de Melo, Aline Carneiro Viana, Marco Fiore, Katia Jaffrès-Runser, Frédéric Le Mouël, Antonio A. F. Loureiro, Lavanya Addepalli, Guangshuo Chen.

Recent studies have analyzed data generated from mobile individuals in urban regions, such as cab drivers or students in large campuses. Particular attention has been paid to the dynamics of user movement, whose real-world complexity cannot be fully captured through synthetic models. Indeed, understanding user mobility is of fundamental importance when designing new communication protocols that exploit opportunistic encounters among users. In this case, the problem mainly lies in correctly forecasting future contacts. To that end, the regularity of daily activities comes in handy, as it enforces periodic (and thus predictable) space-time patterns in human mobility. Although human behavior is characterized by an elevated rate of regularity, random events are always possible in the routines of individuals. Those are hardly predictable situations that deviate from the regular pattern and are unlikely to repeat in the future.

We argue that the ability to accurately spot random and social relationships in dynamic networks is essential to network applications that rely on a precise description of human routines, such as recommendation systems, forwarding strategies and opportunistic dissemination protocols. We thus propose a strategy to analyze users' interactions in mobile networks where users act according to their interests and activity dynamics. Our strategy, named *Random rElationship ClASsifier sTrategy (RECAST)*, allows classifying users' wireless interactions, separating random interactions from different kinds of social ties. To that end, RECAST observes how the real system differs from an equivalent one where entities' decisions are completely random. We evaluate the effectiveness of the RECAST classification on five real-world user contact datasets collected in diverse networking contexts. Our analysis unveils significant differences among the dynamics of users' wireless interactions in the datasets, which we leverage to unveil the impact of social ties on opportunistic routing. We show that, for such specific purpose, the relationships inferred by classifier are more relevant than, e.g., self-declared friendships on Facebook. An initial version of this work was published in the international conference ACM MSWiM 2013 (selected as one of the five better papers of this venue) and an extented version bringing new analysis (e.g., the contact duration-related analysis performed by the internship Lavanya Addepalli and the PhD student Guangshuo Chen) was accepted to be published in 2015 at the Performance Evaluation Elsevier Journal [2].

6.5. Spontaneous Wireless Networks and Internet of Things

internet of things; wireless sensor networks; dissemination; resource management

6.5.1. Network Coding in Large Scale IoT Networks

Participants: Cedric Adjih, Ichrak Amdouni, Hana Baccouch, Antonia Masucci.

We had designed a generic broadcast protocol, called DragonNet, based on network coding and designed for constrained networks such as wireless sensor networks and internet of things. It minimizes the assumptions made of the networks. A variant of this protocol was implemented and run on IoT-LAB: some results were initially presented at IRTF, and a live demo was presented in MASS in october 2014.

6.5.2. Information-Centric Networking in the Internet of Things

Participants: Emmanuel Baccelli, Oliver Hahm, Matthias Waehlisch, Thomas Schmidt, Christian Mehlis.

Within this activity, we explored the feasibility, advantages, and challenges of an ICN-based approach in the Internet of Things. We report on the first NDN experiments in a life-size IoT deployment, spread over tens of rooms on several floors of a building. Based on the insights gained with these experiments, we have analysed the shortcomings of CCN applied to IoT. Several interoperable CCN enhancements are then proposed and evaluated. We significantly decreased control traffic (i.e., interest messages) and leverage data path and caching to match IoT requirements in terms of energy and bandwidth constraints. Our optimizations increase content availability in case of IoT nodes with intermittent activity. Within this activity, we also provided the first experimental comparison of CCN with the common IoT standards 6LoWPAN/RPL/UDP.

IPSO Project-Team

5. New Results

5.1. Highlights of the Year

- E. Faou was plenary speaker at the CANUM, Congrès d'analyse numérique, France, June 2014
- E. Faou was invited to give two presentations in the Analysis and applied mathematics seminars, Cambridge, UK, February 2014.

5.2. Multi-revolution composition methods for highly oscillatory differential equations

In [22], we introduce a new class of multi-revolution composition methods (MRCM) for the approximation of the Nth-iterate of a given near-identity map. When applied to the numerical integration of highly oscillatory systems of differential equations, the technique benefits from the properties of standard composition methods: it is intrinsically geometric and well-suited for Hamiltonian or divergence-free equations for instance. We prove error estimates with error constants that are independent of the oscillatory frequency. Numerical experiments, in particular for the nonlinear Schrödinger equation, illustrate the theoretical results, as well as the efficiency and versatility of the methods.

5.3. Multiscale schemes for the BGK-Vlasov-Poisson system in the quasi-neutral and fluid limits. Stability analysis and first order schemes

In [51], in collaboration with G. Dimarco (University of Ferrara, Italy) and M.-H. Vignal (University of Toulouse), we deal with the development and the analysis of asymptotic stable and consistent schemes in the joint quasi-neutral and fluid limits for the collisional Vlasov-Poisson system. In these limits, the classical explicit schemes suffer from time step restrictions due to the small plasma period and Knudsen number. To solve this problem, we propose a new scheme stable for choices of time steps independent from the small scales dynamics and with comparable computational cost with respect to standard explicit schemes. In addition, this scheme reduces automatically to consistent discretizations of the underlying asymptotic systems. In this first work on this subject, we propose a first order in time scheme and we perform a relative linear stability analysis to deal with such problems. The framework we propose permits to extend this approach to high order schemes in the next future. We finally show the capability of the method in dealing with small scales through numerical experiments.

5.4. Asymptotic preserving scheme for a kinetic model describing incompressible fluids

In [52], in collaboration with M. Lemou (CNRS, Université de Rennes 1) and R. Rao, A. Ruhi, M. Sekhar (Indian Institute of Science, India), the kinetic theory of fluid turbulence modeling developed by Degond and Lemou is considered for further study, analysis and simulation. Starting with the Boltzmann like equation representation for turbulence modeling, a relaxation type collision term is introduced for isotropic turbulence. In order to describe some important turbulence phenomenology, the relaxation time incorporates a dependency on the turbulent microscopic energy and this makes difficult the construction of efficient numerical methods. To investigate this problem, we focus here on a multi-dimensional prototype model and first propose an appropriate change of frame that makes the numerical study simpler. Then, a numerical strategy to tackle the stiff relaxation source term is introduced in the spirit of Asymptotic Preserving Schemes. Numerical tests are performed in a one-dimensional framework on the basis of the developed strategy to confirm its efficiency.

5.5. Comparison of numerical solvers for anisotropic diffusion equations arising in plasma physics

In [39], in collaboration G. Latu (IRFM, Cadarache), we performed a comparison of numerical schemes to approximate anisotropic diffusion problems arising in tokamak plasma physics. We focus on the spatial approximation by using finite volume method and on the time discretization. This latter point is delicate since the use of explicit integrators leads to a severe restriction on the time step. Then, implicit and semi-implicit schemes are coupled to finite volumes space discretization and are compared for some classical problems relevant for magnetically confined plasmas. It appears that the semi-implicit approaches (using ARK methods or directional splitting) turn out to be the most efficient on the numerical results, especially when nonlinear problems are studied on refined meshes, using high order methods in space.

5.6. Asymptotic-Preserving scheme based on a Finite Volume/Particle-In-Cell coupling for Boltzmann- BGK-like equations in the diffusion scaling

In [38], in collaboration with A. Crestetto (University of Nantes), we are concerned with the numerical simulation of the collisional Vlasov equation in the diffusion limit using particles. To that purpose, we use a micro-macro decomposition technique introduced by Bennoune, Lemou and Mieussens. Whereas a uniform grid was used to approximate both the micro and the macro part of the full distribution function in their article, we use here a particle approximation for the kinetic (micro) part, the fluid (macro) part being always discretized by standard finite volume schemes. There are many advantages in doing so: (i) the so-obtained scheme presents a much less level of noise compared to the standard particle method; (ii) the computational cost of the micro-macro model is reduced in the diffusion limit since a small number of particles is needed for the micro part; (iii) the scheme is asymptotic preserving in the sense that it is consistent with the kinetic equation in the rarefied regime and it degenerates into a uniformly (with respect to the Knudsen number) consistent (and deterministic) approximation of the limiting equation in the diffusion regime.

5.7. Hamiltonian splitting for the Vlasov-Maxwell equations

In [23], in collaboration with L. Einkemmer (University of Innsbruck), a new splitting is proposed for solving the Vlasov-Maxwell system. This splitting is based on a decomposition of the Hamiltonian of the Vlasov-Maxwell system and allows for the construction of arbitrary high order methods by composition (independent of the specific deterministic method used for the discretization of the phase space). Moreover, we show that for a spectral method in space this scheme satisfies Poisson's equation without explicitly solving it. Finally, we present some examples in the context of the time evolution of an electromagnetic plasma instability which emphasizes the excellent behavior of the new splitting compared to methods from the literature.

5.8. A hybrid transport-diffusion model for radiative transfer in absorbing and scattering media

In [35], in collaboration with M. Roger (University of Lyon), C. Caliot (CNRS) and P. Coelho (Instituto Superior Tecnico of Lisboa), a new multi-scale hybrid transport-diffusion model for radiative transfer calculations is proposed. In this model, the radiative intensity is decomposed into a macroscopic component calculated by the diffusion equation, and a mesoscopic component. The transport equation for the mesoscopic component allows to correct the estimation of the diffusion equation, and then to obtain the solution of the linear radiative transfer equation. In this work, results are presented for stationary and transient radiative transfer cases, in examples which concern solar concentrated and optical tomography applications. The Monte Carlo and the discrete-ordinate methods are used to solve the mesoscopic equation. It is shown that the multi-scale model allows to improve the efficiency of the calculations when the medium is close to the diffusive regime. Moreover, the development of methods for coupling the radiative transfer equation with the diffusion equation becomes easier with this model than with the usual domain decomposition methods.

5.9. Charge conserving grid based methods for the Vlasov-Maxwell equations

In [26], in collaboration with P. Navaro (CNRS, Strasbourg) and E. Sonnendrücker (IPP Garching, Germany), In this article we introduce numerical schemes for the Vlasov-Maxwell equations relying on different kinds of grid based Vlasov solvers, as opposite to PIC schemes, that enforce a discrete continuity equation. The idea underlying this schemes relies on a time splitting scheme between configuration space and velocity space for the Vlasov equation and on the computation of the discrete current in a form that is compatible with the discrete Maxwell solver.

5.10. Improving conservation properties of a 5D gyrokinetic semi-Lagrangian code

In [32], in collaboration with G. Latu, V. Grandgirard, J. Abiteboul, G. Dif-Pradalier, X. Garbet, P. Ghendrih Y. Sarazin (IRFM, Cadarache), M. Mehrenberger (University of Strasbourg) and E. Sonnendrücker (IPP Garching, Germany), we are concerned with gyrokinetic turbulent simulations, where the knowledge of some stationary states can help reducing numerical artifacts. Considering long-term simulations, the qualities of the Vlasov solver and of the radial boundary conditions have an impact on the conservation properties. In order to improve mass and energy conservation mainly, the following methods are investigated: fix the radial boundary conditions on a stationary state, use a 4D advection operator that avoids a directional splitting, interpolate with a delta-f approach. The combination of these techniques in the semi-Lagrangian code gysela leads to a net improvement of the conservation properties in 5D simulations.

5.11. Simulations of Kinetic Electrostatic Electron Nonlinear (KEEN) Waves with Variable Velocity Resolution Grids and High-Order Time-Splitting

In [16], in collaboration with B. Afeyan (Polymath Research, USA), F. Casa (University Jaume, Spain), A. Dodhy, E. Sonnendrücker (IPP Garching, Germany) and M. Mehrenberger (University of Strasbourg), we are concerned with KEEN waves which are non-stationary, nonlinear, self-organized asymptotic states in Vlasov plasmas. They lie outside the precepts of linear theory or perturbative analysis, unlike electron plasma waves or ion acoustic waves. Steady state, nonlinear constructs such as BGK modes also do not apply. The range in velocity that is strongly perturbed by KEEN waves depends on the amplitude and duration of the ponderomotive force generated by two crossing laser beams, for instance, used to drive them. Smaller amplitude drives manage to devolve into multiple highly-localized vorticlets, after the drive is turned off, and may eventually succeed to coalesce into KEEN waves. Fragmentation once the drive stops, and potential eventual remerger, is a hallmark of the weakly driven cases. A fully formed (more strongly driven) KEEN wave has one dominant vortical core. But it also involves fine scale complex dynamics due to shedding and merging of smaller vortical structures with the main one. Shedding and merging of vorticlets are involved in either case, but at different rates and with different relative importance. The narrow velocity range in which one must maintain sufficient resolution in the weakly driven cases, challenges

fixed velocity grid numerical schemes. What is needed is the capability of resolving locally in velocity while maintaining a coarse grid outside the highly perturbed region of phase space. We here report on a new Semi-Lagrangian Vlasov-Poisson solver based on conservative non-uniform cubic splines in velocity that tackles this problem head on. An additional feature of our approach is the use of a new high-order time-splitting scheme which allows much longer simulations per computational effort. This is needed for low amplitude runs. There, global coherent structures take a long time to set up, such as KEEN waves, if they do so at all. The new code's performance is compared to uniform grid simulations and the advantages are quantified. The birth pains associated with weakly driven KEEN waves are captured in these simulations. Canonical KEEN waves with ample drive are also treated using these advanced techniques. They will allow the efficient simulation of KEEN waves in multiple dimensions, which will be tackled next, as well as generalizations to Vlasov-Maxwell codes. These are essential for pursuing the impact of KEEN waves in high energy density plasmas and in inertial confinement fusion applications. More generally, one needs a fully-adaptive grid-in-phase-space method which could handle all small vorticlet dynamics whether pealing off or remerging. Such fully adaptive grids would have to be computed sparsely in order to be viable. This two-velocity grid method is a concrete and fruitful step in that direction.

5.12. Gyroaverage operator on polar mesh

In [36], in collaboration with C. Steiner, M. Mehrenberger (University of Strasbourg) V. Grandgirard, G. Latu (IRFM, Cadarache). In this work, we are concerned with numerical approximation of the gyroaverage operators arising in plasma physics to take into account the effects of the finite Larmor radius corrections. The work initiated in [Crouseilles, Mehrenberger, Sellama, CiCP 2010] is extended here to polar geometries. A direct method is proposed in the space configuration which consists in integrating on the gyrocircles using interpolation operator (Hermite or cubic splines). Numerical comparisons with a standard method based on a Pade approximation are performed: (i) with analytical solutions, (ii) considering the 4D drift-kinetic model with one Larmor radius and (iii) on the classical linear DIII-D benchmark case [6]. In particular, we show that in the context of a drift-kinetic simulation, the proposed method has similar computational cost as the standard method and its precision is independent of the radius.

5.13. A new fully two-dimensional conservative semi-Lagrangian method: applications on polar grids, from diocotron instability to ITG turbulence

In [25], in collaboration with P. Glanc, S. Hirstoaga, E. Madaule, M. Mehrenberger, J. Pétri (University of Strasbourg), While developing a new semi-Lagrangian solver, the gap between a linear Landau run in 1dx1d and a 5D gyrokinetic simulation in toroidal geometry is quite huge. Intermediate test cases are welcome for checking the code. We consider here as building block, a 2D guiding-center type equation on an annulus. We first revisit a 2D test case previously done with a PIC approach and detail the boundary conditions. We then consider a 4D drift-kinetic slab simulation for which we give some first results of a new conservative method.

5.14. Uniformly accurate numerical schemes for highly oscillatory Klein-Gordon and nonlinear Schrödinger equations

In [21], we are interested in the numerical simulation of nonlinear Schrödinger and Klein-Gordon equations. We present a general strategy to construct numerical schemes which are uniformly accurate with respect to the oscillation frequency. This is a stronger feature than the usual so called "Asymptotic preserving" property, the last being also satisfied by our scheme in the highly oscillatory limit. Our strategy enables to simulate the oscillatory problem without using any mesh or time step refinement, and the orders of our schemes are preserved uniformly in all regimes. In other words, since our numerical method is not based on the derivation and the simulation of asymptotic models, it works in the regime where the solution does not oscillate rapidly, in the highly oscillatory limit regime, and in the intermediate regime with the same order of accuracy. The method is based on two main ingredients. First, we embed our problem in a suitable "two-scale" reformulation with the introduction of an additional variable. Then a link is made with classical strategies based on Chapman-Enskog expansions in kinetic theory despite the dispersive context of the targeted equations, allowing to separate the fast time scale from the slow one. Uniformly accurate (UA) schemes are eventually derived from this new formulation and their properties and performances are assessed both theoretically and numerically.

5.15. Asymptotic preserving schemes for the Wigner-Poisson-BGK equations in the diffusion limit

In [24], we focus on the numerical simulation of the Wigner-Poisson-BGK equation in the diffusion asymptotics. Our strategy is based on a "micro-macro" decomposition, which leads to a system of equations that couple the macroscopic evolution (diffusion) to a microscopic kinetic contribution for the fluctuations. A semi-implicit discretization provides a numerical scheme which is stable with respect to the small parameter ε (mean free path) and which possesses the following properties: (i) it enjoys the asymptotic preserving property in the diffusive limit; (ii) it recovers a standard discretization of the Wigner-Poisson equation in the collisionless regime. Numerical experiments confirm the good behaviour of the numerical scheme in both regimes. The case of a spatially dependent $\varepsilon(x)$ is also investigated.

5.16. Models of dark matter halos based on statistical mechanics: II. The fermionic King model

In [49] we study the fermionic King model which may provide a relevant model of dark matter halos. The exclusion constraint can be due to quantum mechanics (for fermions such as massive neutrinos) or to Lynden- Bells statistics (for collisionless systems undergoing violent relaxation). This model has a finite mass. Dwarf and intermediate size halos are degenerate quantum objects stabilized against gravitational collapse by the Pauli exclusion principle. Large halos at sufficiently high energies are in a gaseous phase where quantum effects are negligible. They are stabilized by thermal motion. Below a critical energy Ec they undergo gravitational collapse (gravothermal catastrophe). This may lead to the formation of a central black hole that does not affect the structure of the halo. This may also lead to the formation of a compact degenerate object surrounded by a hot massive atmosphere extending at large distances. We argue that large dark matter halos should not contain a degenerate nucleus (fermion ball) because these nucleus-halo structures are thermodynamically unstable. We compare the rotation curves of the classical King model to observations of large dark matter halos (Burkert profile). Because of collisions and evaporation, the central density increases while the slope of the halo density profile decreases until an in- stability takes place. We find that the observations are compatible with a King profile at, or close to, the point of marginal stability in the micro- canonical ensemble. At that point, the King profile can be fitted by the modified Hubble profile. This is qualitatively similar to the Burkert profile and discrepancies between the King model and the observations are interpreted as a result of incomplete relaxation.

5.17. Models of dark matter halos based on statistical mechanics: I. The classical King model

In [48] we consider the possibility that dark matter halos are described by the Fermi-Dirac distribution at finite temperature. This is the case if dark matter is a self-gravitating quantum gas made of massive neutrinos at statistical equilibrium. This is also the case if dark matter can be treated as a self-gravitating collisionless gas experiencing Lynden-Bell?s type of violent relaxation. In order to avoid the infinite mass problem and carry out a rigorous stability analysis, we consider the fermionic King model. In this paper, we study the non-degenerate limit leading to the classical King model. This model was initially introduced to describe globular clusters and we propose to apply it also to large dark matter halos where quantum effects are negligible. We study the thermodynamical stability of the different configurations and compare the prediction of the classical King model to the observations of large dark matter halos. Because of collisions and evaporation, the central density increases while the slope of the halo density profile decreases until an instability takes place. We show that large dark matter halos are relatively well-described by the King model at, or close to, the point of marginal microcanonical stability. At that point, the King model generates a density profile that can be approximated by the modified Hubble profile. This profile has a flat core and decreases as r?3 at large distances, like the observational Burkert profile. For large halos, the flat core is due to finite temperature effects, not to quantum mechanics. We argue that statistical mechanics may provide a good description of dark matter halos and interpret the discrepancies as a result of incomplete relaxation like in the case of stellar systems.

5.18. Analysis of models for quantum transport of electrons in graphene layers

In [28], two mathematical models for the self consistent quantum transport of electrons in a graphene layer are presented are analyzed. We treat two situations. First, when the particles can move in all the plane R^2, the model takes the form of a system of massless Dirac equations coupled together by a selfconsistent potential, which is the trace in the plane of the graphene of the 3D Poisson potential associated to surface densities. Second, we consider a situation where the particles are constrained in a regular bounded domain Ω. In order to take into account Dirichlet boundary conditions which are not compatible with the Dirac Hamiltonian H_0, we propose a different model built on a modified Hamiltonian displaying the same energy band diagram as H_0 near the Dirac points.

5.19. Dimension reduction for anisotropic Bose-Einstein condensates in the strong interaction regime

The work [44] deals with the problem of dimension reduction for the three dimensional Gross-Pitaevskii equation (GPE) describing a Bose-Einstein condensate confined in a strongly anisotropic harmonic trap. Since the gas is assumed to be in a strong interaction regime, we have to analyze two combined singular limits: a semi-classical limit in the transport direction and the strong partial confinement limit in the transversal direction.

5.20. Superconvergence of Strang splitting for NLS in T^d

In [47], we investigate the convergence properties of semi-discretized approximations by Strang splitting method applied to fast-oscillating nonlinear Schrödinger equations. Our main contribution is to show that Strang splitting with constant step-sizes is unexpectedly more accurate by a factor ε as compared to established results when the step-size is chosen as an integer fraction of the period, owing to an averaging effect.

5.21. Strong confinement limit for the nonlinear Schrödinger equation constrained on a curve

The preprint [58] is devoted to the cubic nonlinear Schrödinger equation in a two dimensional waveguide with shrinking cross section of order ε. For a Cauchy data living essentially on the first mode of the transverse Laplacian, we provide a tensorial approximation of the solution ψ^ε in the limit $\varepsilon \to 0$, with an estimate of the approximation error, and derive a limiting nonlinear Schrödinger equation in dimension one with an additional effective potential depending on the curvature.

5.22. The fermionic King model

In [50], we study the fermionic King model which may provide a relevant model of dark matter halos.

5.23. Landau damping in Sobolev spaces for the Vlasov-HMF model

In [56], we consider the Vlasov-HMF (Hamiltonian Mean-Field) model. We consider solutions starting in a small Sobolev neighborhood of a spatially homogeneous state satisfying a linearized stability criterion (Penrose criterion). We prove that these solutions exhibit a scattering behavior to a modified state, which implies a nonlinear Landau damping effect with polynomial rate of damping.

5.24. Collisions of vortex filament pairs

In [18], we consider the problem of collisions of vortex filaments for a model introduced by Klein, Majda and Damodaran, and Zakharov to describe the interaction of almost parallel vortex filaments in three-dimensional fluids. Since the results of Crow examples of collisions are searched as perturbations of antiparallel translating pairs of filaments, with initial perturbations related to the unstable mode of the linearized problem; most results are numerical calculations. In this article we first consider a related model for the evolution of pairs of filaments and we display another type of initial perturbation leading to collision in finite time. Moreover we give numerical evidence that it also leads to collision through the initial model. We finally study the self-similar solutions of the model.

5.25. Asymptotic preserving schemes for the Klein-Gordon equation in the non-relativistic limit regime

In [30], we consider the Klein-Gordon equation in the non-relativistic limit regime, i.e. the speed of light c tending to infinity. We construct an asymptotic expansion for the solution with respect to the small parameter depending on the inverse of the square of the speed of light. As the first terms of this asymptotic can easily be simulated our approach allows us to construct numerical algorithms that are robust with respect to the large parameter c producing high oscillations in the exact solution.

5.26. Analysis of a large number of Markov chains competing for transitions

In [17], we consider the behavior of a stochastic system composed of several identically distributed, but non independent, discrete-time absorbing Markov chains competing at each instant for a transition. The competition consists in determining at each instant, using a given probability distribution, the only Markov chain allowed to make a transition. We analyze the first time at which one of the Markov chains reaches its absorbing state. When the number of Markov chains goes to infinity, we analyze the asymptotic behavior of the system for an arbitrary probability mass function governing the competition. We give conditions for the existence of the asymptotic distribution and we show how these results apply to cluster-based distributed systems when the competition between the Markov chains is handled by using a geometric distribution.

5.27. Coexistence phenomena and global bifurcation structure in a chemostat-like model with species-dependent diffusion rates

In [20], we study the competition of two species for a single resource in a chemostat. In the simplest space-homogeneous situation, it is known that only one species survives, namely the best competitor. In order to exhibit coexistence phenomena, where the two competitors are able to survive, we consider a space dependent situation: we assume that the two species and the resource follow a diffusion process in space, on top of the competition process. Besides, and in order to consider the most general case, we assume each population is associated with a distinct diffusion constant. This is a key difficulty in our analysis: the specific (and classical) case where all diffusion constants are equal, leads to a particular conservation law, which in turn allows to eliminate the resource in the equations, a fact that considerably simplifies the analysis and the qualitative phenomena. Using the global bifurcation theory, we prove that the underlying 2-species, stationary, diffusive, chemostat-like model, does possess coexistence solutions, where both species survive. On top of that, we identify the domain, in the space of the relevant bifurcation parameters, for which the system does have coexistence solutions.

5.28. Global behavior of N competing species with strong diffusion: diffusion leads to exclusion

In [46], we study the following problem. For a large class of models involving several species competing for a single resource in a *homogeneous* environment, it is known that the competitive exclusion principle holds: only one species survives eventually. Various works indicate though that coexistence of many species is possible when the competition occurs in a *heterogeneous* environment. We propose here a spatially heterogeneous system modeling several species competing for a single resource, and migrating in the spatial domain. For this model, it is known, at least in particular cases, that if migrations are *slow* enough, then coexistence occurs. In this paper we show at variance that if the spatial migrations are *fast* enough, then our system can be approximated by a spatially homogeneous system, called aggregated model, which can be explicitly computed, and we show that if the competitive exclusion principle holds for the aggregated model, then it holds as well for the original, spatially heterogeneous model. In other words, we show the persistence of the competitive exclusion principle in the spatially heterogeneous situation when migrations are fast. As a consequence, for fast migrations only one species may survive, namely the best competitor *in average*. We last study which is the best competitor *in average* on some examples, and draw some ecological consequences.

5.29. Randomized Message-Passing Test-and-Set

In [42] we present a solution to the well-known Test&Set operation in an asynchronous system prone to process crashes. Test&Set is a synchronization operation that, when invoked by a set of processes, returns yes to a unique process and returns no to all the others. Recently many advances in implementing Test&Set objects have been achieved, however all of them target the shared memory model. In this paper we propose an implementation of a Test&Set object in the message passing model. This implementation can be invoked by any number $p \leq n$ of processes where n is the total number of processes in the system. It has an expected individual step complexity in O(log p) against an oblivious adversary, and an expected individual message complexity in O(n). The proposed Test&Set object is built atop a new basic building block, called selector, that allows to select a winning group among two groups of processes. We propose a message-passing implementation of the selector whose step complexity is constant. We are not aware of any other implementation of the Test&Set operation in the message passing model.

5.30. Existence of densities for the 3D Navier–Stokes equations driven by Gaussian noise

In [27], we prove three results on the existence of densities for the laws of finite dimensional functionals of the solutions of the stochastic Navier-Stokes equations in dimension 3. In particular, under very mild assumptions on the noise, we prove that finite dimensional projections of the solutions have densities with respect to the Lebesgue measure which have some smoothness when measured in a Besov space. This is proved thanks to a new argument inspired by an idea introduced by N. Fournier and J. Printems.

5.31. Diffusion limit for the radiative transfer equation perturbed by a Markovian process

In [54], we study the stochastic diffusive limit of a kinetic radiative transfer equation, which is non linear, involving a small parameter and perturbed by a smooth random term. Under an appropriate scaling for the small parameter, using a generalization of the perturbed test-functions method, we show the convergence in law to a stochastic non linear fluid limit.

5.32. Diffusion limit for the radiative transfer equation perturbed by a Wiener process

In [55], we consider the rigorous derivation of a stochastic non-linear diffusion equation from a radiative transfer equation perturbed with a random noise of white noise type. The proof of the convergence relies on a formal Hilbert expansion and the estimation of the remainder. The Hilbert expansion has to be done up to order 3 to overcome some difficulties caused by the random noise.

KALIFFE Project-Team

6. New Results

6.1. Mixed semi-Lagrangian/finite difference methods for plasma simulations

We present an efficient algorithm for the long time behavior of plasma simulations. We will focus on 4D drift-kinetic model, where the plasma's motion occurs in the plane perpendicular to the magnetic field and can be governed by the 2D guiding-center model.

Hermite WENO reconstructions are applied for solving the Vlasov equation. Here we consider an arbitrary computational domain with an appropriate numerical method for the treatment of boundary conditions.

Then we apply this algorithm for plasma turbulence simulations. We first solve the 2D guiding-center model in a D-shape domain and investigate the numerical stability of the steady state. Then, the 4D drift-kinetic model is studied with a mixed method, *i.e.* the semi-Lagrangian method in linear phase and finite difference method during the nonlinear phase. Numerical results show that the mixed method is efficient and accurate in linear phase and it is much stable during the nonlinear phase. Moreover, in practice it has better conservation properties [9].

6.2. High order semi implicit schemes for PDEs

We consider a new formulation of implicit-explicit (IMEX) methods for the numerical discretization of time dependent partial differential equations. We construct several semi-implicit Runge-Kutta methods up to order three. This method is particularly well suited for problems where the stiff and non-stiff components cannot be well separated. We present different numerical simulations for reaction-diffusion, convection diffusion and nonlinear diffusion system of equations. Finally, we conclude by a stability analysis of the schemes for linear problems.

6.3. A Hierarchy of Hybrid Numerical Methods for Multi-Scale Kinetic Equations

We construct a hierarchy of hybrid numerical methods for multi-scale kinetic equations based on moment realizability matrices, a concept introduced by Levermore, Morokoff and Nadiga. Following such a criterion, one can consider hybrid scheme where the hydrodynamic part is given either by the compressible Euler or Navier-Stokes equations, or even with more general models, such as the Burnett or super-Burnett systems [8].

6.4. Derivation of high order absorbing boundary conditions for the Helmholtz equation in 2D

We present high order absorbing boundary conditions (ABC)for the Helmholtz equation in 2D, that can adapt to any regular shapedsurfaces. The new ABCs are derived by using the technique ofmicro-diagonalisation to approximate the Dirichlet-to-Neumann map.Numerical results on different shapes illustrate the behavior of thenew ABCs along with high-order finite elements [5].

KerData Project-Team

6. New Results

6.1. Highlights of the Year

IEEE Cluster 2014. The KerData Team had a leading role the organization of the IEEE Cluster 2014 conference, held in Madrid (22–26 September 2014): Gabriel Antoniu as PC Chair, Luc Bougé as Student Mentoring Program Chair, Alexandru Costan as Submissions Chair.

6.2. Data Management for Geographically Distributed Workflows

6.2.1. OverFlow: a multi-site-aware framework for Big Data management

Participants: Radu Tudoran, Alexandru Costan, Gabriel Antoniu.

The global deployment of cloud datacenters is enabling large-scale scientific workflows to improve performance and deliver fast responses. This unprecedented geographical distribution of the computation coincides with an increase in the scale of the data handled by such applications, bringing new challenges related to the efficient data management across sites. High throughput, low latencies or cost-related trade-offs are just a few concerns for both cloud providers and users when it comes to handling data across datacenters, as shown in earlier evaluations [21]. Existing solutions are limited to cloud-provided storage, which offers low performance based on rigid cost schemes. In turn, workflow engines need to find ad-hoc substitutes, achieving performance at the cost of complex system configurations, maintenance overheads, reduced reliability and reusability.

We tackle these problems by trying to understand to what extent the intra- and inter-datacenter transfers can impact the total makespan of cloud workflows. We advocate storing data on the compute nodes and transferring files between them directly, in order to exploit data locality and to avoid the overhead of interacting with a shared file system. Under these circumstances, we propose a file management service that enables high throughput through self-adaptive selection among multiple transfer strategies (e.g. FTP-based, BitTorrent-based, etc.). Next, we focus on the more general case of large-scale data dissemination across geographically distributed sites. The key idea is to predict I/O and transfer performance accurately and robustly in a dynamic cloud environment in order to decide judiciously how to perform transfer optimizations over federated datacenters: predict the best combination of protocol and transfer parameters (e.g., multi-routes, flow count, multicast enhancement, replication degree) to maximize throughput or minimize costs, according to users policies. We have implemented these principles in OverFlow, as part of the Azure Cloud so that applications could use it using a Software-as-a-Service (SaaS) approach.

OverFlow [20] was validated on the Microsoft cloud across the 6 EU and US sites. The experiments were conducted on hundreds of nodes using synthetic benchmarks and real-life bio-informatics applications (A-Brain, BLAST). The results show that our system is able to model the cloud performance accurately and to leverage this for efficient data dissemination, being able to reduce the monetary costs and transfer time by up to 3 times.

6.2.2. Metadata management for geographically distributed workflows

Participants: Luis Eduardo Pineda Morales, Radu Tudoran, Alexandru Costan, Gabriel Antoniu.

Scientific workflow data can reach sizes that exceed single-site capabilities. It is needed to support fine-grain data stripping to handle either very large files or very large sets of small files across data centers. Therefore, metadata becomes a critical issue. Moreover, workflow metadata provides crucial information to optimize data management, particularly in the context of geographically distributed data centers. Many present-day distributed file systems, such as GoogleFS and HDFS, include a potential bottleneck as the number of files grows, because they use a centralized metadata management scheme. Thus, we argue for a new, *cloud-based, distributed metadata management* scheme.

We have designed four different approaches to a geographically distributed metadata registry, namely: a) baseline centralized version; b) distributed on each data center with centralized replication agent; c) decentralized non-replicated; and d) decentralized replicated with hierarchical access. A comparative analysis showed that the later strategy performs best in terms of metadata operations per time unit. We then evaluate each of our approaches against various workflow benchmarks, with the purpose of dynamically adapt the metadata handling scheme according to the underlying application and cloud contexts. In the next phase, we will provide a uniform metadata handling tool for scientific workflow engines across cloud datacenters, as well as derive a cost model to offer users the best trade-off (performance vs. cost) driven by their constraints.

6.2.3. Transfer-as-a-Service: a cost-effective model for multi-site cloud data management

Participants: Radu Tudoran, Alexandru Costan, Gabriel Antoniu.

Existing cloud data management solutions are limited to cloud-provided storage, which offers low performance based on rigid cost schemas. Users are therefore forced to design and deploy custom solutions, achieving performance at the cost of complex system configurations, maintenance overheads, reduced reliability and reusability. In [19] we have proposed a dedicated cloud data-transfer service that supports largescale data dissemination across geographically distributed sites, advocating for a Transfer-as-a-Service (TaaS) paradigm. The system aggregates the available bandwidth by enabling multi-route transfers across cloud sites, based on the approach previously described.

We argue that the adoption of such a TaaS approach brings several benefits for both users and the cloud providers who propose it. For users of multi-site or federated clouds, our proposal is able to decrease the variability of transfers and increase the throughput up to three times compared to baseline user options, while benefiting from the well-known high availability of cloud-provided services. For cloud providers, such a service can decrease the energy consumption within a datacenter down to half compared to user-based transfers. Finally, we propose a dynamic cost model schema for the service usage, which enables the cloud providers to regulate and encourage data exchanges via a data transfer market.

6.3. Optimizing Map-Reduce processing

6.3.1. Optimizing Map-Reduce in virtualized environments

Participant: Shadi Ibrahim.

As data-intensive applications become popular in the cloud, their performance on the virtualized platform calls for empirical evaluations and technical innovations. Virtualization has become a prominent tool in data centers and is extensively leveraged in cloud environments: it enables multiple virtual machines (VMs) — with multiple operating systems and applications — to run within a physical server. However, virtualization introduces the challenging issue of providing effective QoS to VMs and preserving the high disk utilization (i.e., reducing the seek delay and rotation overhead) when allocating disk resources to VMs.

In [32], we developed a novel disk I/O scheduling framework, named *Pregather*, to improve disk I/O efficiency through exposure and exploitation of the spatial locality in the virtualized environment (regional and sub-regional spatial locality corresponds to the virtual disk space and applications' access patterns, respectively). In [14], we extend *Pregather* to improve disk I/O utilization further while reducing the disk resource contention and ensuring the I/O performance of VMs with different degrees of spatial locality. To do so, we developed an adaptive time-slice allocation scheme based on the spatial locality of VMs, to adjust the lengths of I/O time slices of VMs dynamically. We evaluated *Pregather* through extensive experiments that involve multiple simultaneous applications of both synthetic benchmarks and a Map-Reduce application (e.g., distributed sort) on Xen-based platforms.

Our evaluations use synthetic benchmarks, a Map-Reduce application (distributed sort) and database workloads. They demonstrate that *Pregather* achieves high disk spatial locality, yields a significant improvement in disk throughput, ensures the performance guarantees of VMs, and enables improved Hadoop performance. This work was done in collaboration with Hai Jin, Song Wu and Xiao Ling from Huazhong University of Science and Technology (HUST).

6.3.2. A simulation approach to evaluate Map-Reduce performance under failure

Participants: Tien Dat Phan, Shadi Ibrahim, Gabriel Antoniu, Luc Bougé.

Map-Reduce is emerging as a prominent tool for large-scale data analysis. It is often advocated as an easier-to-use, efficient and reliable replacement for the traditional programming model of moving the data to the computation. The popular open source implementation of Map-Reduce, Hadoop, is now widely used by major companies, including Facebook, Amazon, Last.fm, and the New York Times. Fault tolerance is one of the key features of the Map-Reduce system. Map-Reduce is designed to handle various kind of failures including stop-fail and time failures: Map-Reduce re-executes failed tasks and re-launches another copy of slow tasks. Although many studies have been dedicated to investigate and improve the performance of Map-Reduce, comparatively little attention has been devoted on investigating the performance of Map-Reduce under failures.

In this ongoing work, we investigate how Map-Reduce (i.e., Hadoop) behaves under failures. To do so, we developed *iHadoop*, a Hadoop simulator developed in Java on top of SimGrid. Experimental results demonstrated that *iHadoop* accurately simulates the behavior of Hadoop and therefore can accurately predict the performance of Hadoop when running on large-scale system using the Grid'5000 testbed. In particular, iHadoop can accurately predict the percentage of Map tasks locality, the number of speculative tasks and, more importantly, the overall execution time of Map-Reduce applications under failures.

6.3.3. Waste-Free Preemption Strategy for Hadoop

Participants: Orçun Yildiz, Shadi Ibrahim, Gabriel Antoniu.

Hadoop is widely used in the computer industry because of its scalability, reliability, ease of use, and low cost of implementation. Hadoop hides the complexity of discovery and handling failures from the schedulers, but the burden of failure recovery relies entirely on users, regardless of root causes. We systematically assess this burden through a set of experiments, and argue that more effort to reduce this cost to users is desirable. We also analyze the drawback of current Hadoop mechanism in prioritizing failed tasks. By trying to launch failed tasks as soon as possible regardless of locality, it significantly increases the execution time of jobs with failed tasks, due to two reasons: 1) available slots might not be free up as quickly as expected; and 2) the slots might belong to machines with no data on it, introducing extra cost for data transfer through network, which is normally the most scare resource in nowadays data centers.

In this ongoing work, we introduce a new algorithmic approach called the waste-free preemption. The waste-free preemption saves Hadoop scheduler from solely choosing between kill, which instantly releases the slots but is wasteful, and wait, which does not waste any previous effort but fails for the two above-mentioned reasons. With this new strategy, a preemptive version of Hadoop's default schedulers (FIFO and Fair) has been implemented. The evaluation demonstrates the effectiveness of the new feature by comparing its performance with the traditional Hadoop mechanism.

6.3.4. Optimizing incremental Map-Reduce computations for on-demand data upload

Participants: Stefan Ene, Alexandru Costan, Gabriel Antoniu.

Research on cloud-based Big Data analytics has focused so far on optimizing the performance and cost-effectiveness of the computations, while largely neglecting an important aspect: users need to upload massive datasets on clouds for their computations. In this context, we study the problem of running Map-Reduce applications by considering the simultaneous optimization of performance and cost of both the data upload and its corresponding computation taken together. We analyze the feasibility of incremental Map-Reduce approaches to let the computation progress as much as possible during the data upload by using already transferred data to compute intermediate results.

Current approaches that are either optimized for different purposes, or address the computational problem independent of the data upload. In contrast, to our best knowledge, this is the first approach which simultaneously focuses on both data upload and processing. In this context, we show in [17] that it is not always efficient to attempt to overlap the transfer time with as many incremental computations as possible: a better solution is to wait long enough to fill the computational capacity of the Map-Reduce cluster. Based on this idea, we developed and evaluated a preliminary prototype. To demonstrate the viability of our prototype in real-life, we run extensive experiments in a distributed setting that involves a 11-node large incremental Map-Reduce deployment based on Hourglass. The results show significant benefits for our approach compared with a simple incremental strategy that starts the next incremental job immediately after the previous has finished: the time-to-solution is improved by 1%, the compute time after the data transfer is finished is reduced by up to 40% and the cost is reduced 10 %-44 %. Compared with a serialized strategy that starts the computation only after all data is transferred, the time-to-solution is improved by up to 30 %, the compute time after the upload finished is reduced by up to 60 % and the cost is reduced between 4 % and 23 %.

6.4. Energy-Aware Data Management in the Cloud and Exascale HPC Systems

6.4.1. Energy-efficiency in Hadoop

Participants: Tien Dat Phan, Shadi Ibrahim, Gabriel Antoniu, Luc Bougé.

With increasingly inexpensive cloud storage and increasingly powerful cloud processing, the cloud has rapidly become the environment to store and analyze data. Most of the large-scale data computations in the cloud heavily rely on the Map-Reduce paradigm and its Hadoop implementation. Nevertheless, this exponential growth in popularity has significantly impacted power consumption in cloud infrastructures.

In [18], we focus on Map-Reduce and we investigate the impact of dynamically scaling the frequency of compute nodes on the performance and energy consumption of a Hadoop cluster. To this end, a series of experiments are conducted to explore the implications of Dynamic Voltage Frequency scaling (DVFS) settings on power consumption in Hadoop-clusters. By adapting existing DVFS governors (i.e., *performance*, *power-save*, *on-demand*, *conservative* and *user-space*) in the Hadoop cluster, we observe significant variation in performance and power consumption of the cluster with different applications when applying these governors: the different DVFS settings are only sub-optimal for different Map-Reduce applications. Furthermore, our results reveal that the current CPU governors do not exactly reflect their design goal and may even become ineffective to manage power consumption in Hadoop clusters.

More recently, we extended our work to further illustrate the behavior of different governors, which influence the energy consumption in Hadoop Map-Reduce. We extend our experimental platform from 15 to 40 nodes and we employ two additional benchmarks: K-means and wordcount. Moreover, we investigate preliminary DVFS models that adjust to the various stages of Hadoop applications. We also demonstrate that achieving better energy efficiency in Hadoop cannot be done by tuning the governors parameters, nor through a naive coarse-grained tuning of the CPU frequencies or the governors according the running phase (i.e., map phase or reduce phase). In addition, we provide an extensive discussion of the sensitivity for different parameters employed in *ondemand* and *conservative* governors.

6.4.2. Exploring the impact of dedicated resources on energy consumption in Exascale systems

Participants: Orçun Yildiz, Matthieu Dorier, Shadi Ibrahim, Gabriel Antoniu.

The advent of fast, unprecedentedly scalable, yet energy-hungry Exascale supercomputers poses a major challenge consisting in sustaining a high performance-per-Watt ratio. While much recent work has explored new approaches to I/O management, aiming to reduce the I/O performance bottleneck exhibited by HPC applications (and hence to improve application performance), there is comparatively little work investigating the impact of I/O management approaches on energy consumption.

In [23], we explore how much energy a supercomputer consumes while running scientific simulations when adopting various I/O management approaches. We closely examine three radically different I/O schemes including time partitioning, dedicated cores, and dedicated nodes. We implement the three approaches within the Damaris I/O middleware and perform extensive experiments with one of the target HPC applications of the Blue Waters sustained-Petaflops supercomputer project: the CM1 atmospheric model. The experimental results obtained on the French Grid'5000 platform highlight the differences between these three approaches and illustrate in which way various configurations of the application and of the system can impact performance and energy consumption.

Based on those experimental results, we are working on building a new energy model which can estimate the energy consumptions of various I/O management approaches and help users in selecting the optimal I/O approach to run their application.

6.4.3. Energy impact of data consistency management in the HBase distributed cloud data store

Participants: Álvaro García Recuero, Shadi Ibrahim, Gabriel Antoniu.

Cloud Computing has recently emerged as a key technology providing individuals and companies with access to remote computing and storage infrastructures. In order to achieve high-availability and fault-tolerance, cloud data storage relies on replication. That comes with the issue of consistency among distant replicas so one can always get the most up-to-date values from any of them (*e.g.*, fresh data).

In that context, being able to provide data consistency and continuous availability in the Cloud is yet a non-trivial problem, mainly due to the ever-increasing volume, variety and velocity of data in storage systems. Big data processing engines (e.g., Hadoop, Spark, etc.) as well as modern NoSQL storage back-ends (HBase, Cassandra) have to therefore deal with these high volumes of information at large scale while still providing applications with a consistent and on-time data delivery.

In this work, a set of synthetic workloads from YCSB (Yahoo! Cloud Service Benchmark) was configured to simulate random reads/writes and measure their impact into the overall energy consumption of a well-known distributed data store, HBase. The cluster is comprised of 40 servers and the results have been confirmed with several configurations and runs on the Grid5000 experimental platform. The results indicate that certain write-intensive workloads can be a bottleneck in terms of throughput, further deepening the problem of having an energy-efficient consistency management. Regarding read-intensive workloads, we observe similar patterns but with a very different impact on their energy footprint. We plan to further investigate how to leverage energy-aware mechanisms that overcome the energy-consistency trade-off, while taking into account the selected configuration.

6.5. Scalable I/O and Visualization for Exascale Systems

6.5.1. CALCioM: mitigating cross-application I/O interference

Participants: Matthieu Dorier, Shadi Ibrahim, Gabriel Antoniu.

As larger supercomputers are used by an increasing number of applications in a concurrent manner, the interference produced by multiple applications accessing a shared parallel file system in contention becomes a major problem. Interference often breaks single-application I/O optimizations (such as access patterns preliminarily optimized to improve data locality on disks), thereby dramatically degrading application I/O performance, increasing run-time variability and, as a result, lowering machine-wide efficiency. We addressed this challenge by proposing CALCioM [15], a framework that aims to mitigate I/O interference through the dynamic selection of appropriate scheduling policies. CALCioM allows several applications running on a supercomputer to communicate and coordinate their I/O strategy in order to avoid interfering with one another. We examined four I/O strategies that can be accommodated in this framework: serializing, interrupting, interfering and coordinating. Experiments on Argonne's BG/P Surveyor machine and on several clusters of Grid'5000 showed that CALCioM can be used to improve the scheduling strategy efficiently and transparently between several otherwise interfering applications, given specified metrics of machine-wide efficiency. This work led to a publication at the IPDPS 2014 conference.

6.5.2. Omnisc'IO: Predicting the I/O patterns of HPC applications

Participants: Matthieu Dorier, Shadi Ibrahim, Gabriel Antoniu.

Many I/O optimizations including prefetching, caching, and scheduling, have been proposed to improve the performance of the I/O stack. In order to optimize these techniques, modeling and predicting spatial and temporal I/O patterns of HPC applications as they run, have become crucial. In this direction we introduced Omnisc'IO [16], an original approach that aims to make a step forward toward an intelligent I/O management of HPC applications in next-generation, post-Petascale supercomputers. It builds a grammar-based model of the I/O behavior of any HPC application, and uses this model to predict when future I/O operations will occur, as well as where and how much data will be accessed. Omnisc'IO is transparently integrated into the POSIX and MPI-I/O stacks and does not require any modification to application sources or to high-level I/O libraries. It works without prior knowledge of the application, and converges to accurate predictions within a couple of iterations only. Its implementation is efficient both in computation time and in memory footprint. Omnisc'IO was evaluated with four real HPC applications — CM1, Nek5000, GTC, and LAMMPS — using a variety of I/O backends ranging from simple POSIX to Parallel HDF5 on top of MPI-I/O. Our experiments showed that Omnisc'IO achieves from 79 % to 100 % accuracy in spatial prediction and an average precision of temporal predictions ranging from 0.2 seconds to less than a millisecond. This work was published at the SC14 conference and initiated the development of the Omnisc'IO software.

6.5.3. Smart In-Situ Visualization

Participants: Lokman Rahmani, Matthieu Dorier, Gabriel Antoniu.

The increasing gap between computational power and I/O performance in new supercomputers has started to drive a shift from an offline approach to data analysis to an inline approach, termed *in-situ visualization* (ISV). While most visualization software now provides ISV, they typically visualize large dumps of unstructured data, by rendering everything at the highest possible resolution. This often negatively impacts the performance of simulations that support ISV, in particular when ISV is performed interactively, as in-situ visualization requires synchronization with the simulation. In this ongoing work, we investigate a smarter method of performing ISV. Our approach consists in adapting the resolution of regions of the visualization area based on how much their data are *relevant* with regards to the physical phenomena being simulated. In this direction, we first provide a generic definition of relevant data subsets based on *data variability*. Following this definition, we investigate various filtering algorithms to detect relevant data subsets automatically. The proposed filtering algorithms are derived from information theory, statistics and image processing. Our work is validated in the context of climate simulation, where we show an up to 40% improvement of time-to-solution without any significant loss regarding the quality of visualization (QoV). QoV loss is *quantified* using the structural similarity index metric (SSIM) that takes in consideration human visual system to compute visual errors.

6.6. Data Streaming and Small Data

6.6.1. JetStream: enabling high-performance event streaming across cloud data-centers

Participants: Radu Tudoran, Alexandru Costan, Gabriel Antoniu.

The easily-accessible computation power offered by cloud infrastructures coupled with the revolution of Big Data are expanding the scale and speed at which data analysis is performed. In their quest for extracting value out of the 3 Vs of Big Data, applications process larger data sets, within and across clouds. Enabling fast data transfers across geographically distributed sites becomes particularly important for applications which manage continuous streams of events in real time. Scientific applications (e.g. the Ocean Observatory Initiative or the ATLAS experiment) as well as commercial ones (e.g. Microsoft's Bing and Office 365 large-scale services) operate on tens of data-centers around the globe and follow similar patterns: they aggregate monitoring data, assess the QoS or run global data mining queries based on inter-site event stream processing.

In [22] we propose a set of strategies for efficient transfers of events between cloud data-centers and we introduce JetStream: a prototype implementing these strategies as a high-performance, batch-based streaming middleware. JetStream is able to self-adapt to the streaming conditions by modeling and monitoring a set of context parameters. It further aggregates the available bandwidth by enabling multi-route streaming across cloud sites. The prototype was validated on tens of nodes from US and Europe data-centers of the Windows

Azure cloud using synthetic benchmarks and with application code in the context of the Alice experiment at CERN. The results show an increase in transfer rate of 250 times over individual event streaming. Besides, introducing an adaptive transfer strategy brings an additional 25 % gain. Finally, the transfer rate can further be tripled thanks to the use of multi-route streaming.

6.6.2. Efficient management of many small data objects

Participants: Pierre Matri, Alexandru Costan, Gabriel Antoniu.

Large-scale intensive applications must often manage millions or even billions of small objects. Twitter, for example, has to record on average 5700 new tweets every second. Each of these objects are typically smaller than a kilobyte, and as a result, the database has to store billions of these objects. The sheer amount of objects and the small data sizes can also be found in many other applications, like sensor networks, or graph processing. Another important aspect are the access patterns of these applications where reads dominate over writes, which means the storage system has to be heavily optimized towards read performance.

To address these challenges, we are designing a novel storage system offering fast data access with minimal overhead. Learning from BlobSeer [33], we introduce a more efficient way to manage metadata. To this end, we propose to remove the centralised version manager and to distribute versions across the whole cluster using a distributed hash table. This greatly reduces the response times by allowing single-hop reads for most usage patterns. Additionally, this approach distributes the load over the whole cluster, thus providing a better horizontal scalability and fault tolerance.

LAGADIC Project-Team

6. New Results

6.1. Visual servoing

6.1.1. Photometric moment-based visual servoing

Participants: Manikandan Bakthavatchalam, François Chaumette.

The goal of this work is to determine an adequate set of visual features to control the six degrees of freedom of a dynamic system. Thanks to a collaboration with Omar Tahri from Le2I in Le Creusot, we have been able to improve the results obtained previously with shifted moments for increasing the stability domain of visual servoing [24].

6.1.2. Histogram-based visual servoing

Participants: Quentin Bateux, Eric Marchand.

Classically visual servoing considered the regulation in the image of a set of visual features (usually geometric features). Recently direct visual servoing schemes, such as photometric visual servoing, have been introduced in order to consider the image as a whole and thus avoid the extraction and the tracking of such geometric features. In this preliminary work, we propose a method to extend direct visual servoing approaches by using a global descriptor, namely intensity histograms, on the whole or multiple sub-sets of the images in order to achieve the control of a 6 degrees of freedom (DoF) robot.

6.1.3. Predictive visual servoing

Participants: Nicolas Cazy, Paolo Robuffo Giordano, François Chaumette.

This study is devoted to the application of predictive control to visual servoing. In a first step, we have developed and compared several predictive models that can be useful when some visual features leave the camera field of view or are lost because of occlusions [25].

6.1.4. Visual servoing of cable-driven parallel robot

Participant: François Chaumette.

This study is realized in collaboration with Rémy Ramadour and Jean-Pierre Merlet from Coprin group at Inria Sophia Antipolis. Its goal is to adapt visual servoing techniques for cable-driven parallel robot in order to achieve accurate manipulation tasks [46]. This study is in the scope of the Inria large-scale initiative action PAL (see Section 8.2.6).

6.1.5. Nanomanipulation

Participants: Le Cui, Eric Marchand.

We began a work, within the ANR P2N Nanorobust project (see Section 8.2.1), on the development of micro- and nano-manipulation within SEM (Scanning Electron Microscope). Our goal is to provide visual servoing techniques for positioning and manipulation tasks with a nanometer precision. This year, we focused on the characterisation of the projection model of a SEM along with the approach required for its calibration [26]. We then address the problem of 6 dof control using photometric feature under an optical microscope [27]. Finally, we focused on the definition of control law able to control the motion along the Z axes with a SEM microscope. Indeed, considering that a SEM is subject to parallel projection model, motion along this axis is not observable. We then address this problem using defocus information. An autofocus process has also been studied.

6.1.6. Audio-based control

Participants: Aly Magassouba, François Chaumette.

This study is not concerned with visual servoing, but to the application of the same principle of closed loop control schemes to audio sensors. It is made in collaboration with Nancy Bertin from Panama group at Inria Rennes-Bretagne Atlantique. In a first step, we have determined the analytical form of the interaction matrix of audio features based on the time difference of arrival on two microphones. From this modeling step, we have determined the different virtual linkages that can be realized in function of the number and configuration of sources.

6.2. Visual navigation of mobile robots

6.2.1. Automous navigation of wheelchairs

Participants: Vishnu Karakkat Narayanan, François Pasteau, Marie Babel, François Chaumette.

The goal of this work is to design an autonomous navigation framework of a wheelchair by means of a single camera and visual servoing. We first focused on a corridor following task where no prior knowledge of the environment is required. The servoing process matches the non-holonomic constraints of the wheelchair and relies on two visual features, namely the vanishing point location and the orientation of the median line formed by the straight lines related to the bottom of the walls. This overcomes the initialization issue typically raised in the literature [9]. The control scheme has been implemented onto a robotized wheelchair and results show that it can follow a corridor with an accuracy of ± 3 cm [16].

We then focused on a door passing task [33]. This doorway passing (and corridor turning) task employs the position of a single doorpost in the image as an input to a Lyapunov-based control scheme which allows the wheelchair to take up a desired trajectory about the doorpost. This trajectory avoids collision with the wall and guarantees that the wheelchair positions itself in front of the doorway regardless of its initial position. Results in simulation demonstrate the convergence and robustness of both control schemes. Experiments conducted on a wheelchair indicate the validity of applying the proposed low-level control system [16].

This study is in the scope of the Inria large-scale initiative action PAL (see Section 8.2.6) as well as of the Apash project (see Section 8.1.1).

6.2.2. Semi-autonomous control of a wheelchair for navigation assistance along corridors

Participants: Vishnu Karakkat Narayanan, Marie Babel, François Pasteau, Alexandre Krupa.

This study concerns a semi-autonomous control approach that we designed for safe wheelchair navigation. The control relies on the combination of primary tasks of wall avoidance as well as door passing performed by a dedicated visual servoing framework and a manual steering task. A smooth transition from manual driving to assisted navigation is obtained thanks to a gradual visual servoing activation method that guarantees the continuity of the control law. The visual servoing task is then progressively activated, when the wheelchair gets closer to the walls or doorposts, in order to avoid collisions [43]. Experimental results clearly show the ability of the approach to provide an efficient solution for wall avoiding and doorway passing purposes [58]. This study is in the scope of the Inria large-scale initiative action PAL (see Section 8.2.6) as well as of the Apash project (see Section 8.1.1). Tests with disabled patients in the rehabilitation center Pôle Saint Hélier (Rennes) are under progress and first results prove the ability of our system to smoothly correct the trajectory of the wheelchair in case of hazardous situations.

Current research works are oriented towards multimodal sensor-based servoing, as well as haptic feedback that leads to an intuitive assistive wheelchair navigation. This work is realized in collaboration with Maud Marchal (Hybrid team). In addition, we are currently working with e-Motion team to design a vision-based human-aware semi-autonomous navigation system.

6.2.3. Social Spacing and human-robot interaction

Participants: Panagiotis Papadakis, Patrick Rives.

A novel probabilistic framework was introduced capable of instantiating diverse models of social spacing and accounting for distinctive dimensions in human-robot interaction, namely, perception capacity and certainty [42]. We have concretely shown how our method allows smooth adaptation in the situation awareness of a robot within common human-robot interaction examples and further showed its utility at the level of path planning by adapting trajectories to social sensitivity levels.

This approach is currently extended to take into account human activity cues in order to enhance robot mapping and navigation and in particular in filtering noisy human detections, detecting passages such as doors and staircases, inferring space occupancy and allowing navigation within unexplored areas.

6.2.4. Target tracking

Participants: Ivan Markovic, François Chaumette.

This study was realized in the scope of the FP7 Regpot Across project (see Section 8.3.1.2) during the three-month visit of Ivan Markovic, Ph.D. student at the University of Zagreb. It consisted in developing a pedestrian visual tracking from an omni-directional fish-eye camera and a visual servoing control scheme so that a mobile robot is able to follow the pedestrian [37]. This study has been validated on our Pioneer robot (see Section 5.4.2).

6.2.5. Obstacle avoidance

Participants: Suman Raj Bista, Fabien Spindler, François Chaumette.

This study was realized in collaboration with Andrea Cherubini who is now Assistant Prof. at Université de Montpellier. It is concerned with our long term researches about visual navigation from a visual memory without any accurate 3D localization [4]. In order to deal with obstacle avoidance while preserving the visibility in the visual memory, we have proposed a control scheme based on tentacles for fusing the data provided by a pan-tilt camera and a laser range sensor [11]. A new study devoted to indoors navigation from segments has started recently.

6.3. Visual tracking and state estimation

6.3.1. 3D model-based tracking

Participant: Eric Marchand.

This study focused on the issue of estimating the complete 3D pose of the camera with respect to a potentially textureless object, through model-based tracking. We proposed to robustly combine complementary geometrical and color edge-based features in the minimization process, and to integrate a multiple-hypotheses framework in the geometrical edge-based registration phase [45]. This method will be tested in the scope of the FP7 RemoveDebris project [36].

6.3.2. Pose estimation through plane tracking

Participants: Aurélien Yol, Eric Marchand.

We proposed a method for localizing an Unmanned Aerial Vehicle (UAV) using georeferenced aerial images. Here we provide a multiple usage localization algorithm based on vision only. To ensure robustness, we choose to use the Mutual Information (MI) within a dense tracking process. MI proved to be very robust toward local and global scene variations. However, dense approaches are often related to drift disadvantages. We solve this problem by using georeferenced images. The localization algorithm has been demonstrated through the localization of a hexarotor UAV fitted with a downward looking camera during real flight tests [53].

6.3.3. 3D tracking of deformable objects

Participants: Bertrand Delabarre, Eric Marchand.

We consider the problem of dense non-rigid visual tracking robust towards global illumination perturbations of the observed scene. The similarity function is based on the sum of conditional variance (SCV). With respect to most approaches that minimize the sum of squared differences, which is poorly robust towards illumination variations in the scene, the choice of SCV as our registration function allows the approach to be naturally robust towards global perturbations. Moreover, a thin-plate spline warping function is considered in order to take into account deformations of the observed template [28].

6.3.4. Structure from motion

Participants: Riccardo Spica, Paolo Robuffo Giordano, François Chaumette.

Structure from motion (SfM) is a classical and well-studied problem in computer and robot vision, and many solutions have been proposed to treat it as a recursive filtering/estimation task. However, the issue of *actively* optimizing the transient response of the SfM estimation error has not received a comparable attention. In the work [18], we showed how to design an online active SfM scheme characterized by an error transient response equivalent to that of a reference linear second-order system with desired poles. Indeed, in a nonlinear context, the observability properties of the states under consideration are not (in general) time-invariant but may depend on the current state and on the current inputs applied to the system. It is then possible to simultaneously act on the estimation gains and system inputs (i.e., the camera velocity for SfM) in order to optimize the observation process and impose a desired transient response to the estimation error. The theory has a general validity and can be applied to many different contexts such as when dealing with point features [18], solid objects like spheres or cylinders [51], or planar regions [47]. Furthermore, the active SfM scheme can also be embedded within a classical visual servoing law exploiting the redundancy of the camera motion w.r.t. the considered visual task [52].

6.3.5. Robust visual odometry

Participants: Tawsif Gokhool, Patrick Rives, Renato José Martins.

Our aim is concentrated around building ego-centric topometric maps represented as a graph of salient keyframe nodes [14]. Additionally, visual odometry from frame to keyframe alignment helps significantly in drift reduction. On the other hand, the sparsity in this kind of graphical representation leads to reduced overlapping between keyframes which can degrade localisation robustness. Our chosen spherical 360^0 field of view (FOV) configuration alleviates the overlapping issue by providing an enriched model of the environment with photometric and geometric information content. Following a multitude of advantages with information fusion, merging of frames in a single representation deals with the problem of data redundancy and sensor noise suppression.

Therefore, the second fold of this work consisted in improving the identified conceptual loopholes above by first proposing a generic uncertainty propagation model as applied to our spherical RGB-D database. Secondly, a probabilistic framework was derived which led to a Mahalanobis inconsistency test incorporating both geometric and photometric uncertainty models [32]. Our framework was further improved by adding up a probabilistic model to filter out dynamic points temporally. Finally, the entire probabilistic framework was applied in order to track the most stable points over time.

6.4. 3D Scene Mapping

6.4.1. New RGB-D sensor design for indoor 3D mapping

Participants: Eduardo Fernandez Moral, Patrick Rives.

A multi-sensor device has been developed for omnidirectional RGB-D (color+depth) image acquisition (see Fig. 3 .a). This device allows to acquire such omnidirectional images at high frame rates (30 Hz). This approach has advantages over other alternatives used today in terms of accuracy and real-time spherical image construction for indoor environments, which are specially interesting for mobile robotics. This device has important prospective applications as fast 3D-reconstruction or Slam.

A calibration method for such device was developed [31], which takes into account the bias of each sensor independently. The proposed calibration method does not require any specific calibration pattern, taking into account the planar structure from the scene to cope with the fact that there is no overlapping between sensors.

In a first instance, this sensor has been exploited for localization and mapping research with mobile robots. For that, the sensor is mounted on a mobile platform together with an standard computer (see Fig. 3 .a). A method to perform image registration and visual odometry has been developed. This method relies in the matching of planar primitives that can be efficiently obtained from the depth images. This technique performs considerably faster than previous registration approaches like ICP, or dense photoconsistency alignment. These last achieve however a better accuracy than our method, what suggests that our method can be used as an initial step to speed-up those.

Slam is also addressed with this device. A solution to this problem using our omnidirectional RGB-D sensor is being researched. The ongoing experiments have shown some initial results for metric-topological pose-graph Slam, where the map consists of a set of spherical keyframes, which are located in a topological arrangement according to their shared observations.

6.4.2. Compact 3D scene representation

Participants: Renato José Martins, Patrick Rives, Tawsif Gokhool.

This work follows in the direction of precise and compact scene representation of large scale environments. The aim is to build a complete geometric and photometric "minimal" model, which is stored within a sparse set of augmented spherical images to asset photo-geometry consistence of the scene from multiple points-of-views. In this direction, an uncertainty model from the full structure combined with those of poses was proposed for point-to-point egocentric fusion. This model allows to reduce sensor noise in a particular keyframe sphere when performing a multi-frame fusion scheme of coherent near information. This first fusion scheme is then improved by exploiting the rigidity/influence of neighboring points representing the surface. For that, an intermediary higher level abstraction of the point cloud is generated by partitioning the input domain into elementary cells, then reducing the number of degrees of freedom and enforcing constraints over the points segmented as being part of the same surface.

The adopted solution is a "weaker" representation of a 3D boundary mesh, based on discontinuous convex planar patches, with the segmentation being done considering the geometry (region growing) or photometry (SLIC superpixels). This synthetic scene built with the planar geometric police proved to well represent the original scene (for both indoor and outdoor real data) with a significant small amount of patches and it is exploited to build robust useful "dynamic" 4D world model, which in turn can be used for assisted/autonomous navigation or virtual reality applications.

6.4.3. Semantic mapping

Participants: Romain Drouilly, Patrick Rives, Panagiotis Papadakis.

Autonomous navigation is one of the most challenging problems to address to allow robots to evolve in our everyday environments. Map-based navigation has been studied for a long time and researches have produced a great variety of approaches to model the world. However, semantic information has only recently been taken into account in those models to improve robot efficiency [56]. The goal of this work is to study how semantics can be used to improve all the steps of navigation process. In a first time, we have developed a new navigation-oriented hybrid metric-topological-semantic model of the world. It captures high-level information and uses it to build extremely compact description of large environments. Then we have used it to design an efficient localization algorithm, able to find a given map content faster than classical methods and allowing human-understandable queries [30]. In a second time, we have studied how semantics can be used to discover unobserved things in the scene. Particularly, we have shown that both statics and dynamic entities, identified by a robot, can inform about the structure of the environment in unobserved areas [29] . We have used this to do "map extrapolation", that is extending a map beyond robot's perceptual limits by reasoning on semantics. This approach has been shown to be of great interest in everyday-life environment. Finally, we have proposed a new scheme for trajectory planing, taking into account not only geometric constraints but also high-level understanding of the world. We have shown the usefulness of this approach to navigate complex environments with highly dynamic areas on both simulated and real-world datasets, well-suited for large outdoor environment navigation.

6.4.4. Augmented reality

Participant: Eric Marchand.

Using Slam methods becomes more and more common in Augmented Reality (AR). To achieve real-time requirement and to cope with scale factor and the lack of absolute positioning issue, we proposed to decouple the localization and the mapping step. This approach has been validated on an Android Smartphone through a collaboration with Orange Labs [38][39]

6.5. Medical robotics

6.5.1. Non-rigid target tracking in ultrasound images based on hierarchical grid interpolation

Participants: Lucas Royer, Jason Chevrie, Marie Babel, Alexandre Krupa.

In order to track the motion of a tumour or cyst during needle insertion, we developed a first approach to track a deformable target within a sequence of 2D ultrasound images. It is based on a dedicated hierarchical grid interpolation algorithm (HGI) that is typically used for real-time video compression purposes. This approach provides a continuous motion representation of the target by using a grid of control points that models both their global displacement and local deformations. The motion of each control point is estimated by a hierarchical and multi-resolution local search method in order to minimize the sum of squared difference of the target pixel intensity between successive images. This approach was validated from 2D ultrasound images of real human tissues undergoing rigid and non-rigid deformations [48] and was recently adapted for tracking 3D deformations.

6.5.2. Non-rigid target tracking in ultrasound images based on physically-based model

Participants: Lucas Royer, Alexandre Krupa.

A second approach for automatically tracking deformable target within 2D ultrasound images has been developed [50]. It combines dense information with a physically-based model and has therefore the advantage of not using any fiducial marker. The physical model is represented by a mass-spring damper system driven by external and internal forces. The external forces are obtained by maximizing an image similarity metric between a reference target and the deformed target along the time. The internal forces of the mass-spring damper system constrain the deformation to be physically plausible and therefore efficiently reduce the sensitivity to the speckle noise. This second approach was validated on simulated and real data, both for rigid and non-rigid motions of soft tissues [49]. It was recently extended for deformable target tracking in 3D ultrasound volumes.

6.5.3. 3D steering of flexible needle by visual servoing

Participants: Alexandre Krupa, Pierre Chatelain.

The objective of this work is to provide robotic assistance during needle insertion procedures such as biopsy or ablation of localized tumor. A method has been developed for steering a beveled-tip flexible needle actuated by a robotic arm in such a way to control the needle curvature in 3D space [34]. It is based on the design of a new duty-cycling control strategy that makes possible to control both the 2 lateral angular velocities and the insertion velocity of the needle tip (3 DOF). An image-based visual servoing approach has then been developed to automatically position the needle tip on a 3D target indicated by the user. It is based on the use of geometrical visual features extracted from 2 images provided by 2 orthogonal cameras observing a translucent gelatin phantom where the needle is inserted. Preliminary results of this automatic targeting task demonstrate the feasibility of this new concept and its robustness to needle kinematic model errors [35]. We recently extended this approach to automatically steer the needle toward a target by an image-based visual servoing that uses geometrical features extracted from images provided by a 3D ultrasound probe.

6.5.4. Optimization of ultrasound image quality by visual servoing

Participants: Pierre Chatelain, Alexandre Krupa.

This study focuses on the automatic positioning of a 2D ultrasound probe in such a way to optimize the quality of the acquired ultrasound images. It is based on the recent framework of ultrasound confidence map, developed in the Chair for Computer Aided Medical Procedures and Augmented Reality of Prof. Nassir Navab, which aims at estimating the per-pixel quality of the ultrasound signal based on a model of sound propagation in soft tissues. In collaboration with Nassir Navab we considered this ultrasound confidence map as a new modality and recently designed a visual servoing control law for image quality optimization.

6.6. Control of single and multiple Unmanned Aerial Vehicles

6.6.1. State estimation and flight control of quadrotor UAVs

Participant: Paolo Robuffo Giordano.

Over the last years the robotics community witnessed an increasing interest in the Unmanned Aerial Vehicle (UAV) field. In particular quadrotor UAVs have become more and more widespread in the community as experimental platform for, e.g., testing novel 3D planning, control and estimation schemes in real-world indoor and outdoor conditions. Indeed, in addition to being able to take-off and land vertically, quadrotors can reach high angular accelerations thanks to the relatively long lever arm between opposing motors. This makes them more agile than most standard helicopters or similar rotorcraft UAVs, and thus very suitable to realize complex tasks such as aerial mapping, air pollution monitoring, traffic management, inspection of damaged buildings and dangerous sites, as well as agricultural applications such as pesticide spraying.

Despite these clear advantages, a clear shortcoming of the quadrotor design lies in its inherent underactuation (only 4 actuated propellers for the 6 dofs of the quadrotor pose). This underactuation limits the quadrotor flying ability in free or cluttered space and, furthermore, it also degrades the possibility of interacting with the environment by exerting desired forces in arbitrary directions. In [17], a novel design for a quadrotor UAV with tilting propellers which is able to overcome these limitations has been presented and experimentally validated. Indeed, the additional set of 4 control inputs actuating the propeller tilting angles can be shown to yield full actuation to the quadrotor position/orientation in space, thus allowing it to behave as a fully-actuated flying vehicle and to overcome the aforementioned underactuation problem.

This work has been realized in collaboration with the Max Planck Institute for Biological Cybernetics, Tübingen, Germany.

6.6.2. Collective control of multiple UAVs

Participant: Paolo Robuffo Giordano.

The challenge of coordinating the actions of multiple robots is inspired by the idea that proper coordination of many simple robots can lead to the fulfilment of arbitrarily complex tasks in a robust (to single robot failures) and highly flexible way. Teams of multi-robots can take advantage of their number to perform, for example, complex manipulation and assembly tasks, or to obtain rich spatial awareness by suitably distributing themselves in the environment. Within the scope of robotics, autonomous search and rescue, firefighting, exploration and intervention in dangerous or inaccessible areas are the most promising applications.

In the context of multi-robot (and multi-UAV) coordinated control, *connectivity* of the underlying graph is perhaps the most fundamental requirement in order to allow a group of robots accomplishing common goals by means of *decentralized* solutions. In fact, graph connectivity ensures the needed continuity in the data flow among all the robots in the group which, over time, makes it possible to share and distribute the needed information. However, connectivity alone is not sufficient to perform certain tasks when only *relative sensing* is used. For these systems, the concept of *rigidity* provides the correct framework for defining an appropriate sensing and communication topology architecture. Rigidity is a combinatorial theory for characterizing the “stiffness” or “flexibility” of structures formed by rigid bodies connected by flexible linkages or hinges. In a broader context, rigidity turns out to be an important architectural property of many multi-agent systems when a common inertial reference frame is unavailable. Applications that rely on sensor fusion for localization, exploration, mapping and cooperative tracking of a target, all can benefit from notions in rigidity theory. The concept of rigidity, therefore, provides the theoretical foundation for approaching decentralized solutions to the aforementioned problems using distance measurement sensors, and thus establishing an appropriate framework for relating system level architectural requirements to the sensing and communication capabilities of the system.

In [22], a decentralized gradient-based rigidity maintenance action for a group of quadrotor UAVs has been proposed and tested in real experimental conditions. By starting in a rigid configuration, the group of UAVs is able to estimate their relative position from sole relative distance measurements, and then use these estimated relative positions in a control action able to preserve rigidity of the whole formation despite presence of sensor limitations (maximum range and line-of-sight occlusions), possible collisions with obstacles and inter-robot collisions. Furthermore, in [54] the rigidity theory has been extended to the case of *bearing measurements*, and directed graphs.

These works were realized in collaboration with the robotics group at the Max Planck Institute for Biological Cybernetics, Tübingen, Germany and with Technion, Israel.

LEAR Project-Team

6. New Results

6.1. Highlights of the Year

- Cordelia Schmid received the **Longuet-Higgins prize** for fundamental contributions in computer vision that have withstood the test of time, 2014.
- We participated to the **Trecvid 2014 Multimedia Event Detection** challenge. We ranked first on one of the four tracks (Ad-hoc training videos with 10 examples per class).
- We participated to the **THUMOS 2014 challenge**. We obtained top ranked results in the localization track of the Thumos 2014 Action Recognition Challenge. The goal of the challenge is to evaluate large-scale action recognition in natural settings.

6.2. Visual recognition in images

6.2.1. Multi-fold MIL Training for Weakly Supervised Object Localization

Participants: Ramazan Cinbis, Cordelia Schmid, Jakob Verbeek.

Object category localization is a challenging problem in computer vision. Standard supervised training requires bounding box annotations of object instances. This time-consuming annotation process is sidestepped in weakly supervised learning. In this case, the supervised information is restricted to binary labels that indicate the absence/presence of object instances in the image, without their locations. In [13], we follow a multiple-instance learning approach that iteratively trains the detector and infers the object locations in the positive training images. Our main contribution is a multi-fold multiple instance learning procedure, which prevents training from prematurely locking onto erroneous object locations. This procedure is particularly important when high-dimensional representations, such as the Fisher vectors, are used. We present a detailed experimental evaluation using the PASCAL VOC 2007 and 2010 datasets. Compared to state-of-the-art weakly supervised detectors, our approach better localizes objects in the training images, which translates into improved detection performance. Figure 1 illustrates the iterative object localization process on several example images.

A journal paper is currently in preparation in which extends [13] by adding experiments with CNN features, and a refinement procedure for the object location inference. These additions improve over related work that has appeared since the publication of the original paper.

6.2.2. Transformation Pursuit for Image Classification

Participants: Mattis Paulin, Jerome Revaud, Zaid Harchaoui, Florent Perronnin [XRCE], Cordelia Schmid.

In this work [19], [23], we use data augmentation (see Fig 2 for examples) to improve image classification performances in a large-scale context. A simple approach to learning invariances in image classification consists in augmenting the training set with transformed versions of the original images. However, given a large set of possible transformations, selecting a compact subset is challenging. Indeed, all transformations are not equally informative and adding uninformative transformations increases training time with no gain in accuracy. We propose a principled algorithm – Image Transformation Pursuit (ITP) – for the automatic selection of a compact set of transformations. ITP works in a greedy fashion, by selecting at each iteration the one that yields the highest accuracy gain. ITP also allows to efficiently explore complex transformations, that combine basic transformations. We report results on two public benchmarks: the CUB dataset of bird images and the ImageNet 2010 challenge. Using Fisher Vector representations, we achieve an improvement from 28.2% to 45.2% in top-1 accuracy on CUB, and an improvement from 70.1% to 74.9% in top-5 accuracy on ImageNet. We also show significant improvements for deep convnet features: from 47.3% to 55.4% on CUB and from 77.9% to 81.4% on ImageNet.

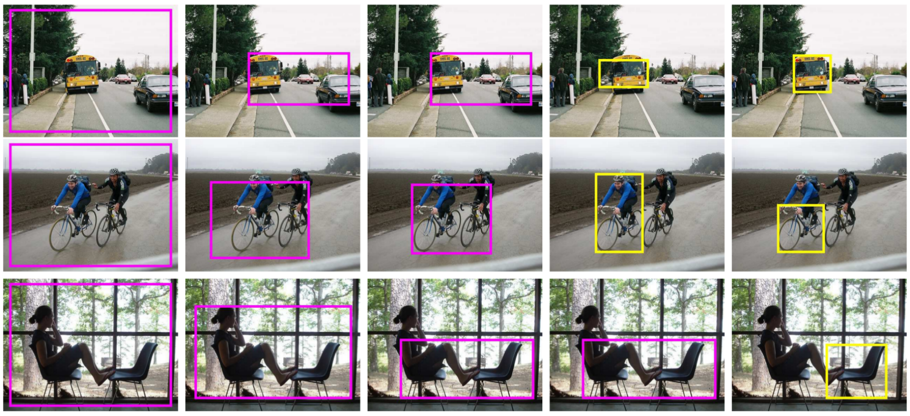


Figure 1. Illustrattion of our iterative object localization process on several example images, from initialization (left) to final localization (right). Yellow bounding boxes indicate that the object location hypothesis is in agreement with the ground-truth, for pink boxes the hypothesis is incorrect.

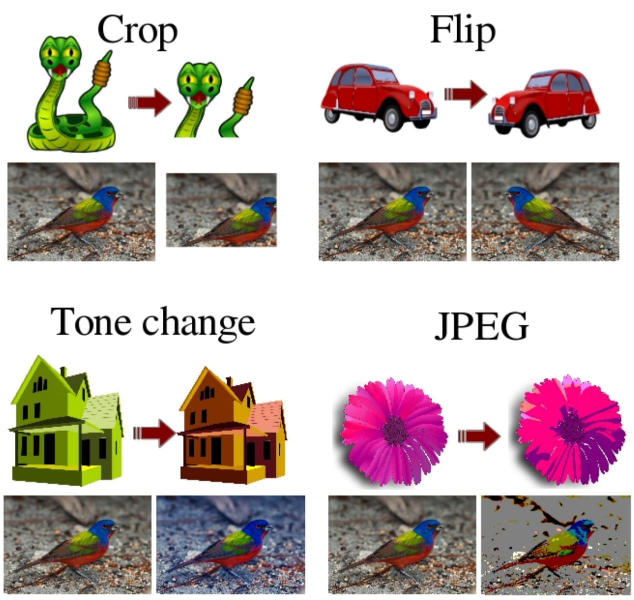
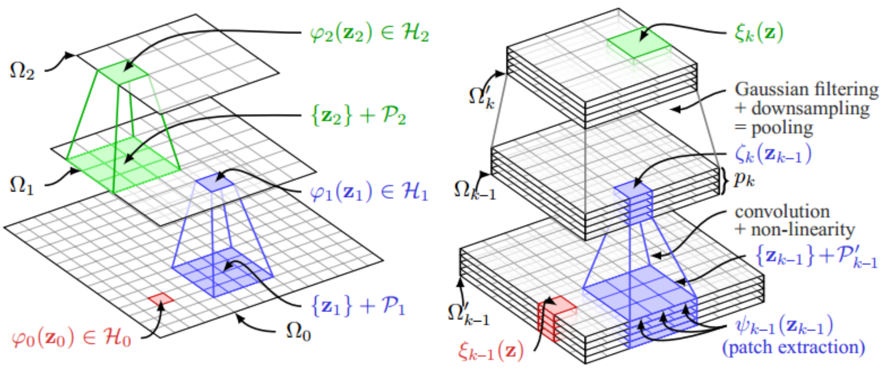


Figure 2. Examples of transformations used in [19], [23].

6.2.3. Convolutional Kernel Networks

Participants: Julien Mairal, Piotr Koniusz, Zaid Harchaoui, Cordelia Schmid.

An important goal in visual recognition is to devise image representations that are invariant to particular transformations. In this paper [16] we address this goal with a new type of convolutional neural network (CNN) whose invariance is encoded by a reproducing kernel. Unlike traditional approaches where neural networks are learned either to represent data or for solving a classification task, our network learns to approximate the kernel feature map on training data. Such an approach enjoys several benefits over classical ones. First, by teaching CNNs to be invariant, we obtain simple network architectures that achieve a similar accuracy to more complex ones, while being easy to train and robust to overfitting. Second, we bridge a gap between the neural network literature and kernels, which are natural tools to model invariance. We evaluate our methodology on visual recognition tasks where CNNs have proven to perform well, e.g., digit recognition with the MNIST dataset, and the more challenging CIFAR-10 and STL-10 datasets, where our accuracy is competitive with the state of the art. Figure 3 illustrates the architecture of our network.



(a) Hierarchy of image feature maps. (b) Zoom between layer k–1 and k of the CKN.

Figure 3. Left: concrete representation of the successive layers for the multilayer convolutional kernel. Right: one layer of the convolutional neural network that approximates the kernel.

6.2.4. Scene Text Recognition and Retrieval for Large Lexicons

Participants: Udit Roy [IIIT Hyderabad, India], Anand Mishra [IIIT Hyderabad, India], Karteek Alahari, C. v. Jawahar [IIIT Hyderabad, India].

In [21], we propose a framework for recognition and retrieval tasks in the context of scene text images. In contrast to many of the recent works, we focus on the case where an image-specific list of words, known as the small lexicon setting, is unavailable. We present a conditional random field model defined on potential character locations and the interactions between them. Observing that the interaction potentials computed in the large lexicon setting are less effective than in the case of a small lexicon, we propose an iterative method, which alternates between finding the most likely solution and refining the interaction potentials. We evaluate our method on public datasets and show that it improves over baseline and state-of-the-art approaches. For example, we obtain nearly 15% improvement in recognition accuracy and precision for our retrieval task over baseline methods on the IIIT-5K word dataset, with a large lexicon containing 0.5 million words.

6.2.5. On Learning to Localize Objects with Minimal Supervision

Participants: Hyun On Song [UC Berkeley], Ross Girschick [UC Berkeley], Stefanie Jegelka [UC Berkeley], Julien Mairal, Zaid Harchaoui, Trevor Darrell [UC Berkeley].

Learning to localize objects with minimal supervision is an important problem in computer vision, since large fully annotated datasets are extremely costly to obtain. In this paper [22], we propose a new method that achieves this goal with only image-level labels of whether the objects are present or not. Our approach combines a discriminative submodular cover problem for automatically discovering a set of positive object windows with a smoothed latent SVM formulation. The latter allows us to leverage efficient quasiNewton optimization techniques. Experimental results are presented in Figure 4 .

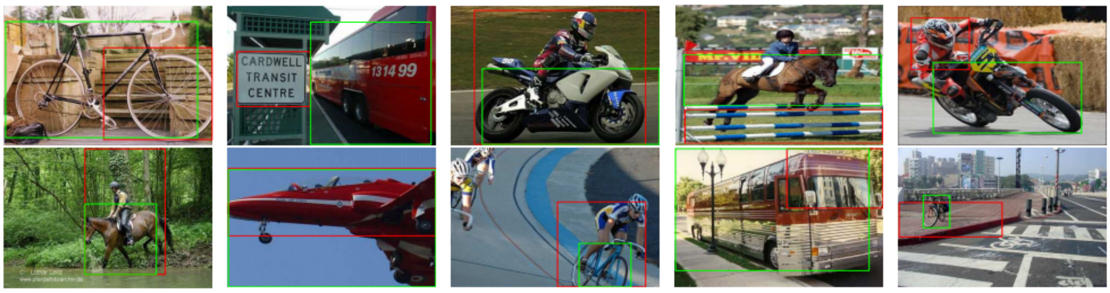


Figure 4. Visualization of some common failure cases of constructed positive windows by (Siva et al., 2012) vs our method. Red bounding boxes are constructed positive windows from (Siva et al., 2012). Green bounding boxes are constructed positive windows from our method.

6.2.6. Good Practice in Large-Scale Learning for Image Classification

Participants: Zeynep Akata, Florent Perronnin [XRCE], Zaid Harchaoui, Cordelia Schmid.

In this paper [3], we benchmark several SVM objective functions for large-scale image classification. We consider one-vs-rest, multi-class, ranking, and weighted approximate ranking SVMs. A comparison of online and batch methods for optimizing the objectives shows that online methods perform as well as batch methods in terms of classification accuracy, but with a significant gain in training speed. Using stochastic gradient descent, we can scale the training to millions of images and thousands of classes. Our experimental evaluation shows that ranking-based algorithms do not outperform the one-vs-rest strategy when a large number of training examples are used. Furthermore, the gap in accuracy between the different algorithms shrinks as the dimension of the features increases. We also show that learning through cross-validation the optimal rebalancing of positive and negative examples can result in a significant improvement for the one-vs-rest strategy. Finally, early stopping can be used as an effective regularization strategy when training with online algorithms. Following these “good practices”, we were able to improve the state-of-the-art on a large subset of 10K classes and 9M images of ImageNet from 16.7% Top-1 accuracy to 19.1%.

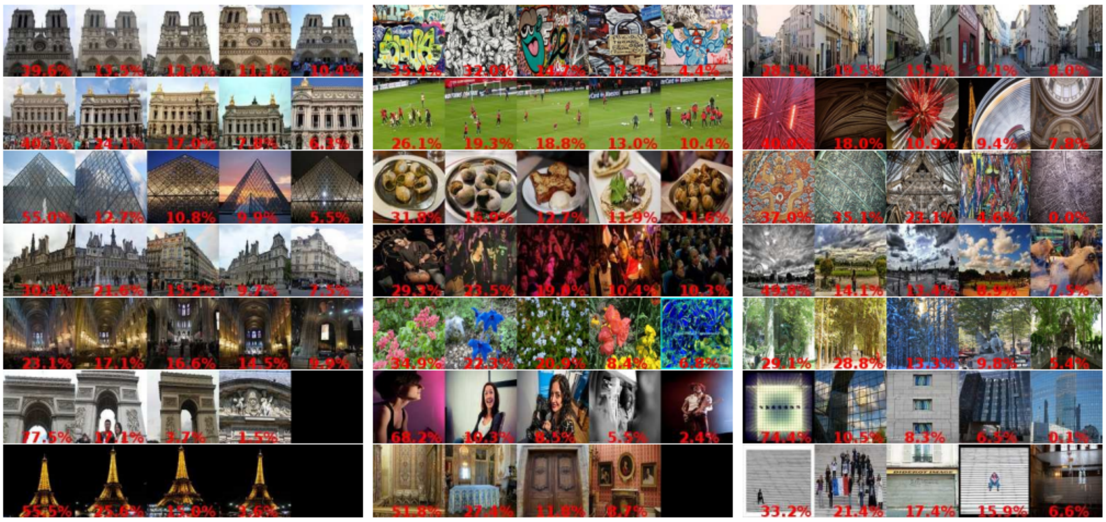
6.3. Learning and statistical models

6.3.1. Fast and Robust Archetypal Analysis for Representation Learning

Participants: Yuansi Chen, Julien Mairal, Zaid Harchaoui.

In [9], we revisit a pioneer unsupervised learning technique called archetypal analysis, which is related to successful data analysis methods such as sparse coding and non-negative matrix factorization. Since it was proposed, archetypal analysis did not gain a lot of popularity even though it produces more interpretable models than other alternatives. Because no efficient implementation has ever been made publicly available, its application to important scientific problems may have been severely limited. Our goal is to bring back into favour archetypal analysis. We propose a fast optimization scheme using an active-set strategy, and provide an efficient open-source implementation interfaced with Matlab, R, and Python. Then, we demonstrate the usefulness of archetypal analysis for computer vision tasks, such as codebook learning, signal classification, and large image collection visualization.

In Figure 5 , we present some archetypes corresponding to the request “Paris” when downloading 36 600 images uploaded in 2012 and 2013, and sorted by relevance on the Flickr website.



(a) Archetypes representing (expected) landmarks. (b) Less expected archetypes. (c) Archetypes representing scene composition.

Figure 5. Classical landmarks appear on the left, which is not surprising since Flickr contains a large number of vacation pictures. In the middle, we display several archetypes that we did not expect, including ones about soccer, graffitis, food, flowers, and social gatherings. Finally, we display on the right some archetypes that do not seem to have some semantic meaning, but they capture some scene composition or texture that are common in the dataset.

6.3.2. Conditional Gradient Algorithms for Norm-Regularized Smooth Convex Optimization

Participants: Zaid Harchaoui, Anatoli Juditsky, Arkadii Nemirovski.

In this paper [6], we consider convex optimization problems arising in machine learning in high-dimensional settings. For several important learning problems, such as e.g. noisy matrix completion, state-of-the-art optimization approaches such as composite minimization algorithms are difficult to apply and do not scale up to large datasets. We study three conditional gradient-type algorithms, *i.e.* first-order optimization algorithms that require a linear minimization oracle but do not require a proximal oracle. These new algorithms are suitable for large-scale problems, and enjoy finite-time convergence guarantees. Promising experimental results are presented on two large-scale real-world datasets. The method is illustrated in Figure 6 .

6.3.3. A Smoothing Approach for Composite Conditional Gradient with Nonsmooth Loss

Participants: Federico Pierucci, Zaid Harchaoui, Jérôme Malick [BIPOP Team, Inria].

In [25], we consider learning problems where the nonsmoothness lies both in the convex empirical risk and in the regularization penalty. Examples of such problems include learning with nonsmooth loss functions and atomic decomposition regularization penalty. Such doubly nonsmooth learning problems prevent the use of recently proposed composite conditional gradient algorithms for training, which are particularly attractive for large-scale applications. Indeed, they rely on the assumption that the empirical risk part of the objective is smooth. We propose a composite conditional gradient algorithm with smoothing to tackle such learning

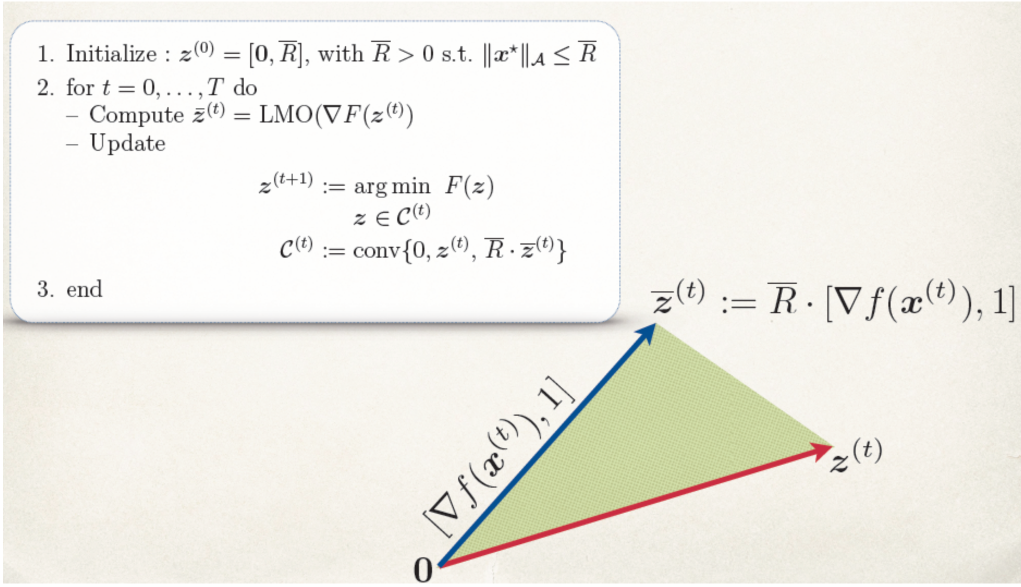


Figure 6. Overview of the composite conditional gradient algorithm which minimizes $F(x) := f(x) + \lambda\|x\|_{\mathcal{A}}$, where f is smooth and $\|\cdot\|_{\mathcal{A}}$ is an atomic-decomposition norm.

problems. We set up a framework allowing to systematically design parametrized smooth surrogates of nonsmooth loss functions. We then propose a smoothed composite conditional gradient algorithm, for which we prove theoretical guarantees on the accuracy. We present promising experimental results on collaborative filtering tasks (see Figure 7).

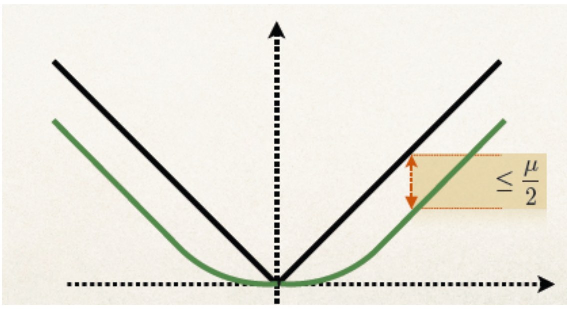


Figure 7. Illustration of the smooth surrogate with parameter μ (green) of the absolute value function (black).

6.3.4. Incremental Majorization-Minimization Optimization with Application to Large-Scale Machine Learning

Participant: Julien Mairal.

In this paper [27], we study optimization methods consisting of iteratively minimizing surrogates of an objective function, as illustrated in Figure 8 . We introduce a new incremental scheme that experimentally matches or outperforms state-of-the-art solvers for large-scale optimization problems typically arising in machine learning.

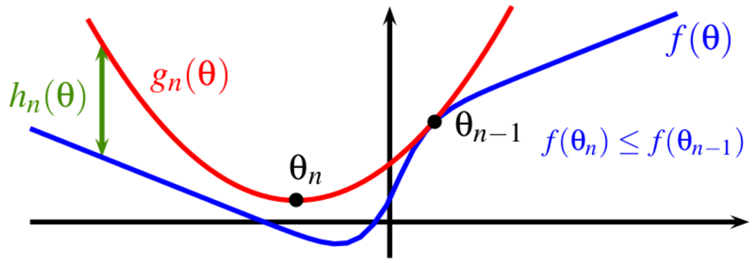


Figure 8. Illustration of the basic majorization-minimization principle. We compute a surrogate g_n of the objective function f around a current estimate θ_{n-1}. The new estimate θ_n is a minimizer of g_n. The approximation error h_n is smooth.

6.3.5. Efficient RNA Isoform Identification and Quantification from RNA-Seq Data with Network Flows

Participants: Elsa Bernard [Institut Curie, Ecoles des Mines-ParisTech], Laurent Jacob [CNRS, LBBE Laboratory], Julien Mairal [correspondant], Jean-Philippe Vert [Institut Curie, Ecoles des Mines-ParisTech].

Several state-of-the-art methods for isoform identification and quantification are based on ℓ_1-regularized regression, such as the Lasso. However, explicitly listing the—possibly exponentially—large set of candidate transcripts is intractable for genes with many exons. For this reason, existing approaches using the ℓ_1-penalty are either restricted to genes with few exons or only run the regression algorithm on a small set of preselected isoforms. In [4], we introduce a new technique called FlipFlop, which can efficiently tackle the sparse estimation problem on the full set of candidate isoforms by using network flow optimization. Our technique removes the need of a preselection step, leading to better isoform identification while keeping a low computational cost. Experiments with synthetic and real RNA-Seq data confirm that our approach is more accurate than alternative methods and one of the fastest available. Figure 9 presents the graph on which the network flow optimization is performed.

6.3.6. Riemannian Sparse Coding for Positive Definite Matrices

Participants: Anoop Cherian, Suvrit Sra [MPI].

Inspired by the great success of sparse coding for vector valued data, our goal in this work [12] is to represent symmetric positive definite (SPD) data matrices as sparse linear combinations of atoms from a dictionary, where each atom itself is an SPD matrix. Since SPD matrices follow a non-Euclidean (in fact a Riemannian) geometry, existing sparse coding techniques for Euclidean data cannot be directly extended. Prior works have approached this problem by defining a sparse coding loss function using either extrinsic similarity measures (such as the log-Euclidean distance) or kernelized variants of statistical measures (such as the Stein divergence, Jeffrey's divergence, etc.). In contrast, we propose to use the intrinsic Riemannian distance on the manifold of SPD matrices. Our main contribution is a novel mathematical model for sparse coding of SPD matrices; we also present a computationally simple algorithm for optimizing our model. Experiments on several computer vision datasets showcase superior classification and retrieval performance compared against state-of-the-art approaches.

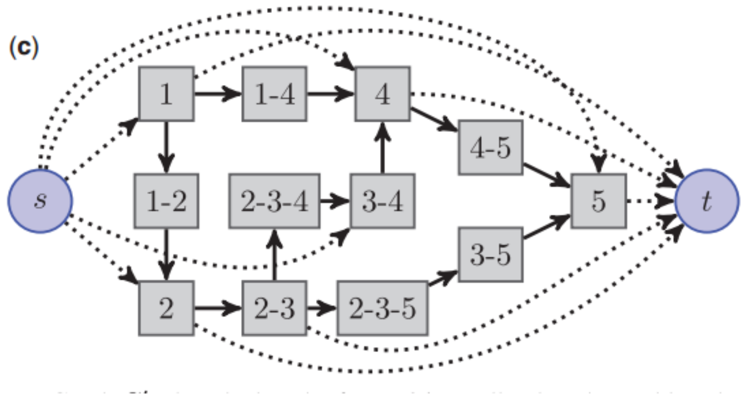


Figure 9. Graph on which we perform network flow optimization. Nodes represent observed reads, and paths on the graph correspond to isoforms.

6.4. Recognition in video

6.4.1. Occlusion and Motion Reasoning for Long-Term Tracking

Participants: Yang Hua, Karteek Alahari, Cordelia Schmid.

Object tracking is a reoccurring problem in computer vision. Tracking-by-detection approaches, in particular Struck, have shown to be competitive in recent evaluations. However, such approaches fail in the presence of long-term occlusions as well as severe viewpoint changes of the object. In this paper we propose a principled way to combine occlusion and motion reasoning with a tracking-by-detection approach. Occlusion and motion reasoning is based on state-of-the-art long-term trajectories which are labeled as object or background tracks with an energy-based formulation. The overlap between labeled tracks and detected regions allows to identify occlusions. The motion changes of the object between consecutive frames can be estimated robustly from the geometric relation between object trajectories. If this geometric change is significant, an additional detector is trained. Experimental results show that our tracker obtains state-of-the-art results and handles occlusion and viewpoints changes better than competing tracking methods. This work corresponds to the publication [15] and is illustrated in Figure 10 .

6.4.2. Category-Specific Video Summarization

Participants: Danila Potapov, Matthijs Douze, Zaid Harchaoui, Cordelia Schmid.

In large video collections with clusters of typical categories, such as “birthday party” or “flash-mob”, category-specific video summarization can produce higher quality video summaries than unsupervised approaches that are blind to the video category. Given a video from a known category, our approach published in [20] first efficiently performs a temporal segmentation into semantically-consistent segments, delimited not only by shot boundaries but also general change points. Then, equipped with an SVM classifier, our approach assigns importance scores to each segment. The resulting video assembles the sequence of segments with the highest scores, as shown in Figure 11 . The obtained video summary is therefore both short and highly informative. Experimental results on videos from the multimedia event detection (MED) dataset of TRECVID’11 show that our approach produces video summaries with higher relevance than the state of the art.

6.4.3. Efficient Action Localization with Approximately Normalized Fisher Vectors

Participants: Dan Oneata, Jakob Verbeek, Cordelia Schmid.

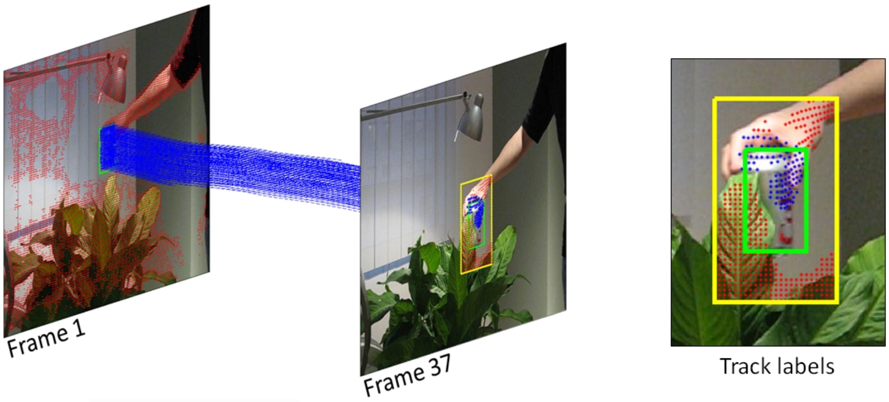


Figure 10. Left: Long-term tracks beginning in frame 1 of the Coke sequence. The yellow box shows the search region used to compute the bounding box most likely to contain the object (green box). We use the tracks to estimate the object state. Right: Close-up of the track labels in frame 37. Here, less than 60% of the tracks within the predicted bounding box are assigned to the object (blue), and the remaining are labelled as background (red). Thus, the object is predicted to be in an occluded state.

Figure 11. Original video, and its video summary for the category "birthday party".

The Fisher vector (FV) representation is a high-dimensional extension of the popular bag-of-word representation. Transformation of the FV by power and ℓ_2 normalizations has shown to significantly improve its performance, and led to state-of-the-art results for a range of image and video classification and retrieval tasks. These normalizations, however, render the representation non-additive over local descriptors. Combined with its high dimensionality, this makes the FV computationally expensive for the purpose of localization tasks. In [18] we present approximations to both these normalizations (see Figure 12), which yield significant improvements in the memory and computational costs of the FV when used for localization. Second, we show how these approximations can be used to define upper-bounds on the score function that can be efficiently evaluated, which enables the use of branch-and-bound search as an alternative to exhaustive sliding window search. We present experimental evaluation results on classification and temporal localization of actions in videos. These show that the our approximations lead to a speedup of at least one order of magnitude, while maintaining state-of-the-art action recognition and localization performance.

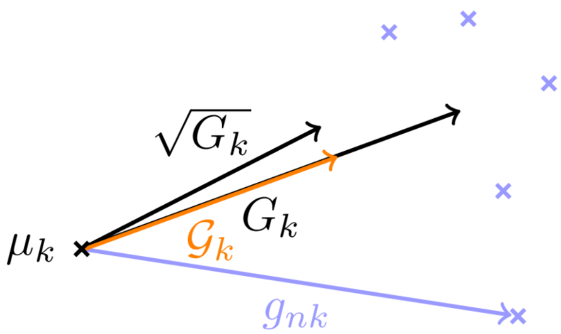


Figure 12. Schematic illustration of the proposed approximation for the square-root normalization. We depict a Fisher vector G_k as an aggregation of individual gradients g_{nk}. Both the exact ($\sqrt{G_k}$) and the approximated ($\mathcal{G}_k$) square-root normalizations scale similarly the Fisher vector G_k; the approximated variant has the property of preserving the orientation of the Fisher vector G_k.

6.4.4. Spatio-Temporal Object Detection Proposals

Participants: Dan Oneata, Jakob Verbeek, Cordelia Schmid, Jerome Revaud.

Spatio-temporal detection of actions and events in video is a challenging problem. Besides the difficulties related to recognition, a major challenge for detection in video is the size of the search space defined by spatio-temporal tubes formed by sequences of bounding boxes along the frames. Recently methods that generate unsupervised detection proposals have proven to be very effective for object detection in still images. These methods open the possibility to use strong but computationally expensive features since only a relatively small number of detection hypotheses need to be assessed. In [17] we make two contributions towards exploiting detection proposals for spatio-temporal detection problems. First, we extend a recent 2D object proposal method, to produce spatio-temporal proposals by a randomized supervoxel merging process (see Figure 13). We introduce spatial, temporal, and spatio-temporal pairwise supervoxel features that are used to guide the merging process. Second, we propose a new efficient supervoxel method. We experimentally evaluate our detection proposals, in combination with our new supervoxel method as well as existing ones. This evaluation shows that our supervoxels lead to more accurate proposals when compared to using existing state-of-the-art supervoxel methods.

6.4.5. EpicFlow: Edge-Preserving Interpolation of Correspondences for Optical Flow

Participants: Revaud Jerome, Weinzaepfel Philippe, Harchaoui Zaid, Cordelia Schmid.

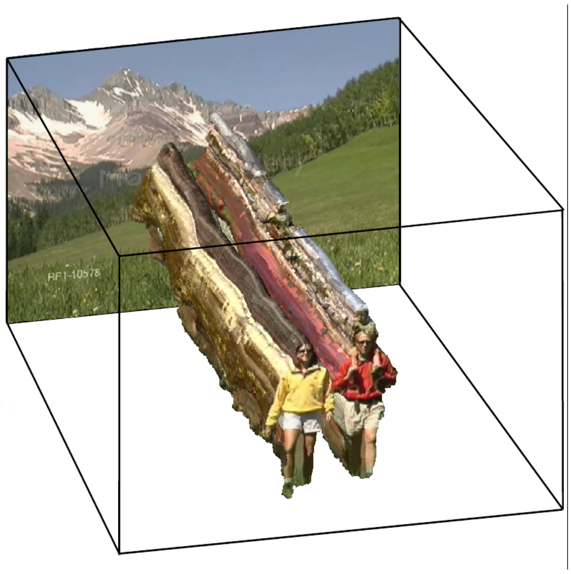


Figure 13. An example of a spatio-temporal proposal generated by our randomized supervoxel merging process. The video sample belongs to the UCF Sports dataset.

We propose a novel approach [29] for optical flow estimation, targeted at large displacements with significant occlusions. It consists of two steps: i) dense matching by edge-preserving interpolation from a sparse set of matches; ii) variational energy minimization initialized with the dense matches. The sparse-to-dense interpolation relies on an appropriate choice of the distance, namely an edge-aware geodesic distance. This distance is tailored to handle occlusions and motion boundaries (see Figure 14), two common and difficult issues for optical flow computation. We also propose an approximation scheme for the geodesic distance to allow fast computation without loss of performance. Subsequent to the dense interpolation step, standard one-level variational energy minimization is carried out on the dense matches to obtain the final flow estimation. The proposed approach, called Edge-Preserving Interpolation of Correspondences (EpicFlow) is fast and robust to large displacements. It significantly outperforms the state of the art on MPI-Sintel and performs on par on KITTI and Middlebury.

6.4.6. Weakly Supervised Action Labeling in Videos Under Ordering Constraints.

Participants: Piotr Bojanowski [Willow team, Inria], Rémi Lajugie [Willow team, Inria], Francis Bach [Sierra team, Inria], Ivan Laptev [Willow team, Inria], Jean Ponce [Willow team, Inria], Cordelia Schmid, Josef Sivic [Willow team, Inria].

Suppose we are given a set of video clips, each one annotated with an ordered list of actions, such as “walk” then “sit” then “answer phone” extracted from, for example, the associated text script. See Fig. 15 for an illustration. In this work [8], we seek to temporally localize the individual actions in each clip as well as to learn a discriminative classifier for each action. We formulate the problem as a weakly supervised temporal assignment with ordering constraints. Each video clip is divided into small time intervals and each time interval of each video clip is assigned one action label, while respecting the order in which the action labels appear in the given annotations. We show that the action label assignment can be determined together with learning a classifier for each action in a discriminative manner. We evaluate the proposed model on a new and challenging dataset of 937 video clips with a total of 787720 frames containing sequences of 16 different actions from 69 Hollywood movies.

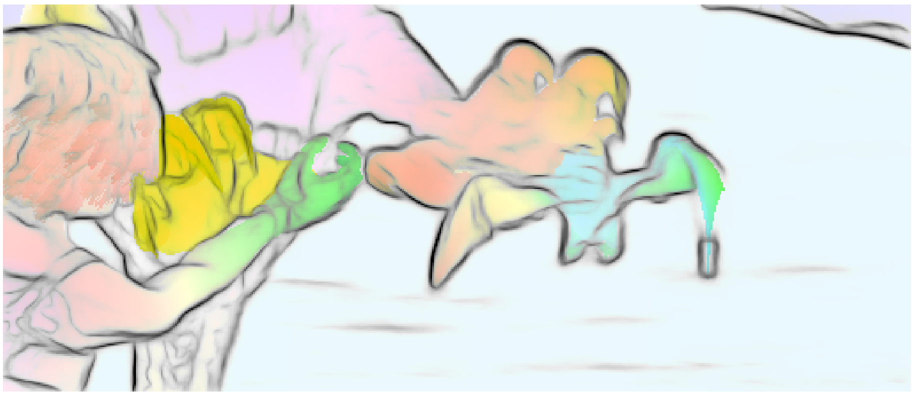


Figure 14. Image edges detected with SED and ground-truth optical flow. Motion discontinuities appear most of the time at image edges

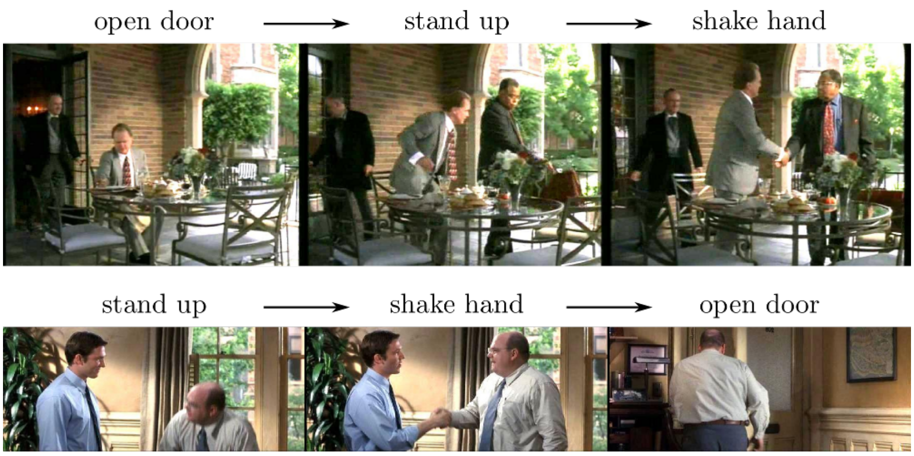


Figure 15. Sample data used as input to our method. Every video clip comes with an ordered list of actions that appears in it. These actions are not temporaly localized, only the order is known. The goal of our paper is to correctly localize these actions according to a discriminative criterion.

6.4.7. Mixing Body-Part Sequences for Human Pose Estimation

Participants: Cherian Anoop, Mairal Julien, Alahari Karteek, Schmid Cordelia.

This work [11] presents a method for estimating articulated human poses in videos. We cast this as an optimization problem defined on body parts with spatio-temporal links between them. The resulting formulation is unfortunately intractable and previous approaches only provide approximate solutions. Although such methods perform well on certain body parts, e.g., head, their performance on lower arms, i.e., elbows and wrists, remains poor. We present a new approximate scheme with two steps dedicated to pose estimation. First, our approach takes into account temporal links with subsequent frames for the less-certain parts, namely elbows and wrists. Second, our method decomposes poses into limbs, generates limb sequences across time, and recomposes poses by mixing these body part sequences (See Figure 16 for an illustration). We introduce a new dataset "Poses in the Wild", which is more challenging than the existing ones, with sequences containing background clutter, occlusions, and severe camera motion. We experimentally compare our method with recent approaches on this new dataset as well as on two other benchmark datasets, and show significant improvement.

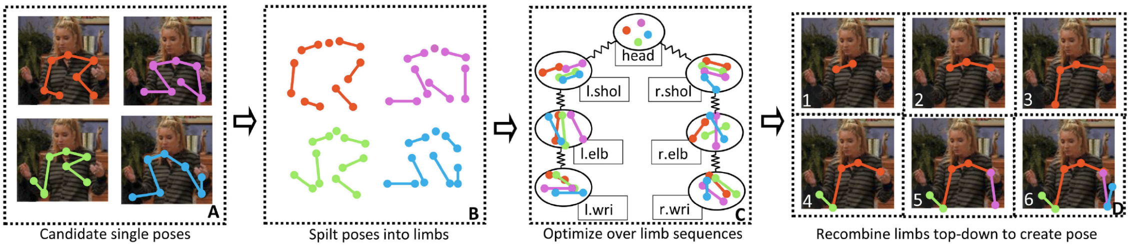


Figure 16. Illustration of our limb recombination scheme. From left to right: Block-A: An image and four candidate poses, where only a part of each pose is well-aligned with the person. Block-B: We divide each candidate pose into limb parts. Block-C: We allow the recombination of limbs from different pose candidates with constraints between two limbs that have a joint in common. Block-D: An example where recombination builds an accurate pose, which is not in the original candidate set.

6.4.8. The LEAR Submission at Thumos 2014

Participants: Dan Oneata, Jakob Verbeek, Cordelia Schmid.

In [28] we describe the submission of our team to the THUMOS workshop in conjunction with ECCV 2014. Our system is based on Fisher vector (FV) encoding of dense trajectory features (DTF), which we also used in our 2013 submission. The dataset is based on the UCF101 dataset, which is currently the largest action dataset both in terms of number of categories and clips, with more than 13000 clips drawn from 101 action classes. This year special attention was paid to classification of uncropped videos, where the action of interest appears in videos that contain also non-relevant sections. This year's submission additionally incorporated static-image features (SIFT, Color, and CNN) and audio features (ASR and MFCC) for the classification task. For the detection task, we combined scores from the classification task with FV-DTF features extracted from video slices. We found that these additional visual and audio feature significantly improve the classification results. For localization we found that using the classification scores as a contextual feature besides local motion features leads to significant improvements. In Figure 17 we show the middle frame from the top four ranked videos corresponding to the three hardest classes (as evaluated on the validation data). Our team has ranked second on the classification challenge (out of eleven teams) and first on the detection challenge (out of three teams).

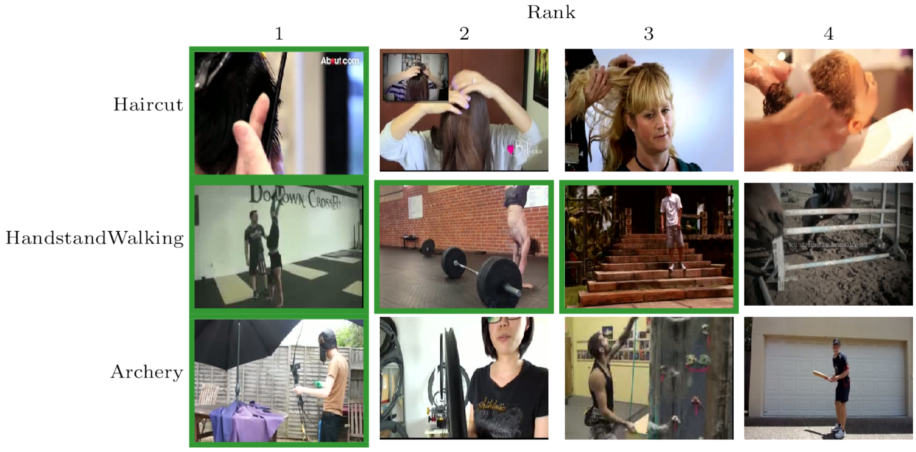


Figure 17. Snapshots from the top four ranked test videos for the three hardest classes; green borders indicate true positives.

6.4.9. The LEAR Submission at TrecVid MED 2014

Participants: Matthijs Douze, Dan Oneata, Mattis Paulin, Clément Leray, Nicolas Chesneau, Danila Potapov, Jakob Verbeek, Karteek Alahari, Zaid Harchaoui, Lori Lamel [Spoken Language Processing group, LIMSI, CNRS], Jean-Luc Gauvain [Spoken Language Processing group, LIMSI, CNRS], Christoph Schmidt [Fraunhofer IAIS, Sankt Augustin], Cordelia Schmid.

In [26] we describe our participation to the 2014 edition of the TrecVid Multimedia Event Detection task. Our system is based on a collection of local visual and audio descriptors, which are aggregated to global descriptors, one for each type of low-level descriptor, using Fisher vectors. Besides these features, we use two features based on convolutional networks: one for the visual channel, and one for the audio channel. Additional high-level features are extracted using ASR and OCR features. Finally, we used mid-level attribute features based on object and action detectors trained on external datasets. In the notebook paper we present an overview of the features and the classification techniques, and experimentally evaluate our system on TrecVid MED 2011 data.

We participated in four tasks, which differ in the amount of training videos for each event (either 10 or 100), and the time that is allowed for the processing. For the 20 pre-specified events several weeks are allowed to extract features, train models, and to score the test videos (which consisted of 8,000 hours of video this year). For the 10 ad-hoc events, we only have five days to do all processing. Across the 11 participating teams, our results ranked first for the 10-example ad-hoc task, and fourth and fifth place for the other tasks.

LEMON Team

6. New Results

6.1. Highlights of the Year

Antoine ROUSSEAU and 5 co-authors released in 2014 the book *Brèves de Maths* [16]. This work (in french) selected more than 100 posts from the blog breves-de-maths.fr, in the framework of the international initiative "Mathematics of the Planet Earth". In this book (see cover 5), no complicated numbers, no weird equation, but short and clear sentences together with nice drawings to illustrate everyday life topics on our planet with the beauty of mathematics.

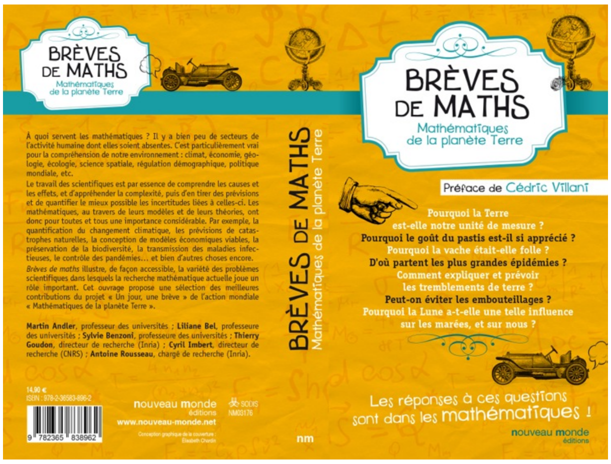


Figure 5. Brèves de Maths. Ed. Nouveau Monde, 2014

6.2. A Schwarz coupling method for dimensionally heterogeneous problem

Participant: Antoine Rousseau.

We study and analyze in [10] an efficient iterative coupling method for a dimensionally heterogeneous problem. We consider the case of 2-D Laplace equation with non symmetric boundary conditions with a corresponding 1-D Laplace equation. We first show how to obtain the 1-D model from the 2-D one by integration along one direction, by analogy with the link between shallow water equations and the Navier-Stokes system. Then, we focus on the design of an Schwarz-like iterative coupling method. We discuss the choice of boundary conditions at coupling interfaces. We prove the convergence of such algorithms and give some theoretical results related to the choice of the location of the coupling interface, and the control of the difference between a global 2-D reference solution and the 2-D coupled one. These theoretical results are illustrated numerically.

6.3. Bioremediation of water ressources

Participants: Antoine Rousseau, Alexis Pacholik.

Together with fellows from the MODEMIC team, we proposed a strategy for the bioremediation of water ressources such as lakes or lagoons. The originality is that the water treatment has to be done outside of the resource, in order not to dislocate its fragile ecological equilibrium.
The objective is to reach a targeted acceptable state for the resource within the minimal time. The patent [19] has been filed in connection with this work.

6.4. A well-balanced and positive preserving DG scheme for the SW equations

Participants: Arnaud Duran, Fabien Marche.

We consider in [5] the discontinuous Galerkin discretization of the nonlinear Shallow Water equations on unstructured triangulations. We propose an efficient combination of ingredients that leads to a simple high-order robust and well-balanced scheme, based on the alternative formulation of the equations known as the pre-balanced shallow water equations. We show that the preservation of the motionless steady states can be achieved, for an arbitrary order of polynomial expansion. Additionally, the preservation of the positivity of the water height is ensured using the recent method introduced in [59]. Some comparisons with a recent finite-volume MUSCL approach are also performed. The well-known tsunami test case shown in figures 6 and 7 has been computed here with high order DG scheme on unstructured triangulation.

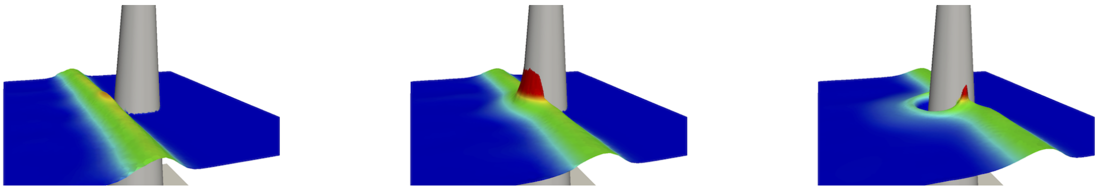


Figure 6. Tsunami wave over a conical island - Lateral view of the free surface at times t=5,6 and 7s.

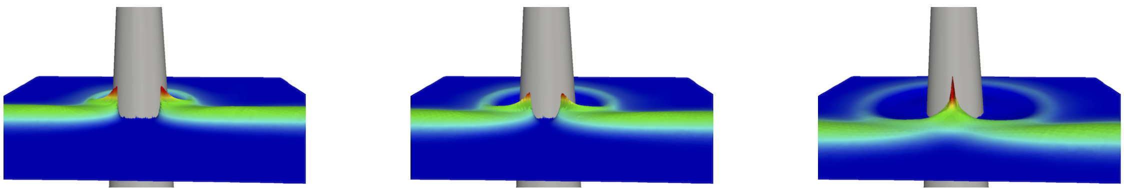


Figure 7. Tsunami wave over a conical island - Rear view of the free surface at times t=8,9 and 10s.

6.5. A well-balanced and positive preserving DG scheme for the GN equations

Participants: Arnaud Duran, Fabien Marche.

We introduce in [4] a discontinuous-Galerkin Finite-Element method to approximate the solutions of a new family of 1d Green-Naghdi models. These new models are shown to be more computationally efficient, while being asymptotically equivalent to the initial formulation with regard to the shallowness parameter. Using the free surface instead of the water height as a conservative variable, the models are recasted under a pre-balanced formulation and discretized using a nodal expansion basis. Independently from the polynomial degree in the approximation space, the preservation of the motionless steady-states is automatically ensured, and the water height positivity is enforced. A simple numerical procedure devoted to stabilize the computations in the vicinity of broken waves is also described. The validity of the resulting model is assessed through extensive numerical validations.

6.6. A new class of fully nonlinear and weakly dispersive Green-Naghdi models

Participant: Fabien Marche.

We introduce in [8] a new class of two-dimensional fully nonlinear and weakly dispersive Green-Naghdi equations over varying topography. These new Green-Naghdi systems share the same order of precision as the standard one but have a mathematical structure which makes them much more suitable for the numerical resolution, in particular in the demanding case of two dimensional surfaces. For these new models, we develop a high order, well balanced, and robust numerical code relying on a hybrid finite volume and finite difference splitting approach. The hyperbolic part of the equations is handled with a high-order finite volume scheme allowing for breaking waves and dry areas. The dispersive part is treated with a finite difference approach. Higher order accuracy in space and time is achieved through WENO reconstruction methods and through an SSP-RK time stepping. Particular effort is made to ensure positivity of the water depth.

6.7. Upscaling transfer properties in heterogeneous porous media

Participant: Vincent Guinot.

In [9] the passive solute transport was studied in a periodic, artificial porous medium. A Laplace analysis of the breakthrough curves indicates that the widely used, classical Advection-Dispersion (AD) model cannot reproduce the contaminant transport features accurately. Neither can fractional dynamics-based, anomalous dispersion models. The models failing to reproduce the features of contaminant transport is shown to be due to the Fick-like, gradient-based operator used to represent dispersion, that induces infinite signal propagation speed, even when fractional models are used. The Laplace analysis shows that advection processes are predominant at all time and space scales. The size of the Representative Elementary Volume is shown to be 20 to 30 periods.

LFANT Project-Team

5. New Results

5.1. Highlights of the Year

Aurel Page has defended his PhD thesis on *Méthodes explicites pour les groupes arithmétiques* [12] in July 2014. Nicolas Mascot has defended his PhD thesis on *Computing modular Galois representations* [11], in July 2014.

5.2. Class groups and other invariants of number fields

Participants: Karim Belabas, Jean-Paul Cerri, Pierre Lezowski.

In [21], P. Lezowski describes the explicit computation of the Euclidean minimum of a number field. It has been published in Mathematics of Computation.

Ohno and Nakagawa have proved, relations between the counting functions of certain cubic fields. These relations may be viewed as complements to the Scholz reflection principle, and Ohno and Nakagawa deduced them as consequences of 'extra functional equations' involving the Shintani zeta functions associated to the prehomogeneous vector space of binary cubic forms. In [26], Henri Cohen, Simon Rubinstein-Salzedo and Frank Thorne generalize their result by proving a similar identity relating certain degree fields with Galois groups D and F respectively, for any odd prime, and in particular we give another proof of the Ohno–Nakagawa relation without appealing to binary cubic forms.

The article [16] by H. Cohen and F. Thorne, H. Cohen on Dirichlet series associated to cubic fields with given resolvent has been published. This article gives an explicit formula for the Dirichlet series $\sum_K |\Delta(K)|^{-s}$, where the sum is over isomorphism classes of all cubic fields whose quadratic resolvent field is isomorphic to a fixed quadratic field k.

This work is extended in [15] where H. Cohen give efficient numerical methods for counting exactly the number of D_ℓ number fields of degree ℓ with given quadratic resolvent, for calculating the constants occurring in their asymptotic expansions, and give tables for typical cases.

5.3. Number and function fields

Participants: Jean-Marc Couveignes, Karim Belabas.

In the article [29], J. Brau study the growth of the Galois invariants of the p-Selmer group of an elliptic curve in a degree p Galois extension. He shows that this growth is determined by certain local cohomology groups and determine necessary and sufficient conditions for these groups to be trivial.

In the article [30] written with J. Nathan, J. Brau study the modular curve $X'(6)$ of level 6 defined over $\mathbb{Q}$ whose $\mathbb{Q}$-rational points correspond to j-invariants of elliptic curves E over $\mathbb{Q}$ for which $\mathbb{Q}(E[2])$ is a subfield of $\mathbb{Q}(E[3])$. They characterize the j-invariants of elliptic curves with this property by exhibiting an explicit model of $X'(6)$. $X'(6)(\mathbb{Q})$ gives an infinite family of examples of elliptic curves with non-abelian "entanglement fields," which is relevant to the systematic study of correction factors of various conjectural constants for elliptic curves over $\mathbb{Q}$.

5.4. Quaternion algebras

Participants: Jean-Paul Cerri, Pierre Lezowski, Aurel Page.

In the article [14] written with J. Chaubert, J.-P. Cerri and P. Lezowski study totally indefinite Euclidean quaternion fields over a number field K, that is to say where no infinite place is ramified. Relying on some generalisation of Hasse–Schilling–Maaß Norm Theorem, they prove that the Euclidean property of K implies the Euclidean property of any totally indefinite Euclidean quaternion field over K. Conversely, they provide the complete list of norm-Euclidean and totally indefinite quaternion fields over an imaginary quadratic number field. In particular, the article exhibits a totally indefinite and norm-Euclidean quaternion field over a non-Euclidean number field. This provides an answer to a question by Eichler. The proofs are both theoretic and algorithmic. The article has been published in Acta Arithmetica.

Deciding whether an ideal of a number field is principal and finding a generator is a fundamental problem with many applications in computational number theory. In the article [25] gives a an algorithm for indefinite quaternion algebras by reducing the decision problem to that in the underlying number field. It also gives an heuristically subexponential algorithm for finding a generator.

5.5. Complex multiplication and modularity

Participants: Jean-Marc Couveignes, Andreas Enge, Nicolas Mascot, Enea Milio, Aurel Page, Damien Robert.

A. Enge and E. Thomé describe in [20] a quasi-linear algorithm for computing Igusa class polynomials of Jacobians of genus 2 curves via complex floating-point approximations of their roots. After providing an explicit treatment of the computations in quartic CM fields and their Galois closures, they pursue an approach due to Dupont for evaluating θ-constants in quasi-linear time using Newton iterations on the Borchardt mean. They report on experiments with the implementation CMH and present an example with class number 20016.

In [34] E. Milio explains how to generalise the work of Régis Dupont for computing modular polynomials in dimension 2 to invariants derived from theta constants. Modular polynomials have many applications. In particular, they could speed up the CRT-algorithm to compute class fields of degree 4 CM-fields which would lead to faster algorithms to construct cryptographically secure Jacobians of hyperelliptic curves. They are also used to compute graphs of isogenies. This paper presents how to compute modular polynomials and the polynomials computed and then it proves some of their properties.

With F. Morain, A. Enge has determined exhaustively under which conditions "generalised Weber functions", that is, simple quotients of η functions of not necessarily prime transformation level and not necessarily of genus 1, yield class invariants [19]. The result is a new infinite family of generators for ring class fields, usable to determine complex multiplication curves. They examine in detail which lower powers of the functions are applicable, thus saving a factor of up to 12 in the size of the class polynomials, and describe the cases in which the polynomials have integral rational instead of integral quadratic coefficients.

N. Mascot has continued his work on computing Galois representations attached to Jacobians of modular curves. He has given tables of modular Galois representations in [33] obtained using the algorithm of [39]. He has computed Galois representations modulo primes up to 31 for the first time. In particular, he has computed the representations attached to a newform with non-rational (but of course algebraic) coefficients, which had never been done before. These computations take place in the Jacobians of modular curves of genus up to 26.

5.6. Elliptic curve and Abelian varieties cryptology

Participants: Jean-Marc Couveignes, Andreas Enge, Damien Robert.

In [27] J.-M. Couveignes and T. Ezome show how to efficiently evaluate functions, including Weil functions and canonical Theta functions, on Jacobian varieties and their quotients. They deduce a quasi-optimal algorithm to compute (l, l) isogenies between Jacobians of genus two curves, using a compact representation and differential characterisation of isogenies in this context. This work has been submitted to the LMS Journal of Computation and Mathematics.

The paper [18] by J.-M. Couveignes and R. Lercier describing the problem of parameterisations by radicals of low genus algebraic curves has been accepted in *Advances in mathematics of communications*.

In [31] D. Lubicz and D. Robert explain how to improve the arithmetic of Abelian and Kummer varieties. The speed of the arithmetic is a crucial factor in the performance of abelian varieties based cryptosystem. Depending on the cryptographic application, the speed record holder are elliptic curves (in the Edwards model) or the Kummer surface of an hyperelliptic curves of genus 2 (in the level 2 theta model). One drawback of the Kummer surface is that only scalar multiplications are available, which may be a drawback in certain cryptographic protocols. The previous known models to work on the Jacobian rather than the Kummer surface (Mumford coordinates or theta model of level 4) are too slow and not competitive with Elliptic Curves. This paper explains how to use geometric properties (like projective normality) to speed up the arithmetic. In particular it introduces a novel addition algorithm on Kummer varieties (compatible additions), and use it to enhance multi-exponentiations in Kummer varieties and to obtain new models of abelian surfaces where the scalar multiplication is as fast as on the Kummer surface.

In [32] (which has been accepted at LMS Journal of Computation and Mathematics), D. Lubicz and D. Robert explain how to compute isogenies between abelian varieties given algebraic equation of the kernel. The previous algorithms to compute isogenies between abelian varieties needed the coordinates of generators of the kernel. One drawback was that even if the kernel is rational, these generators may live in extension of large degree, especially for Abelian varieties defined over a number field rather than a finite field. This paper combines the use of formal coordinates together with a normalisation alongs linear subspaces of the kernel rather than the whole kernel to derive an algorithm which is quasi-optimal if the degree of the isogeny is ℓ^g, for ℓ congruent to 1 modulo 4.

This article expands the article [17] by D. Cosset and D. Robert about the computation of (ℓ, ℓ)-isogenies in dimension 2 which has been published in Mathematics of Computation.

5.7. Pairings

Participants: Andreas Enge, Damien Robert.

The article [22] by D. Lubicz and D. Robert explaining how to compute optimal pairings on abelian varieties described by their theta models has been accepted for publication at Journal of Symbolic Computation.

In [24], A. Enge and J. Milan report on the APIP implementation of cryptographic pairings on elliptic curves in PARI/GP. For security levels equivalent to the different AES flavours, they exhibit suitable curves in parametric families and show that optimal ate and twisted ate pairings exist and can be efficiently evaluated. They provide a correct description of Miller's algorithm for signed binary expansions such as the NAF and extend a recent variant due to Boxall et al. to addition-subtraction chains. They analyse and compare several algorithms proposed in the literature for the final exponentiation. Finally, they give recommendations on which curve and pairing to choose at each security level.

LIFEWARE Team

6. New Results

6.1. Highlight: Xavier Duportet laureate of the AEF docteurs-entrepreneurs prize

Xavier Duportet has been awarded the AEF prize at the docteurs-entrepreneurs competition. His thesis, made jointly within Lifeware and the Weiss lab at MIT, was entitled "Developing new tools and platforms for mammalian synthetic biology: from the assembly and chromosomal integration of large genetic circuits to the engineering of artificial intercellular communication systems". He published his research in *Nucleic Acids Research* and *Nature Biotechnology* [7], [5]. In particular, he demonstrated the assembly and chromosomal integration in mammalian cells of the largest gene circuit integrated to date. Subsequently, he co-founded the startup company PhageX. He was also a laureate of the Concours National de Création d'Entreprises Innovantes and the Concours Mondial d'Innovation (personalized medicine track).

He is the president of the Hello Tomorrow challenge and vice-president of the Osons La France initiative. He has notably been featured in articles published in Le Monde, L'Obs, and L'Opinion. He has been an invited speaker at the prestigious 4th Congreso del Futuro in Santiago (Chili). [7]

6.2. Highlight: François Fages laureate of the French Academy of Sciences

François Fages was very honoured to receive the Michel Monpetit prize 2014 of the French Academy of Sciences for his contributions to fundamental computer science (unification theory and constraint logic programming) and computational systems biology (modeling of biochemical networks and design and supervision of the implementation of the BIOCHAM software).

6.3. Highlight: Pauline Traynard Best Student Paper Prize at CMSB 2014, for Trace Simplifications preserving Temporal Logic Formulae with Case Study in a Coupled Model of the Cell Cycle and the Circadian Clock

Participants: François Fages, Sylvain Soliman, Pauline Traynard.

Pauline Traynard was very pleased to receive the Best Student Paper Prize of the twelth International Conference on Computational Methods for Systems Biology, 17-19 November 2014, Univ. of Manchester, UK, for a communication on trace simplifications preserving temporal logic properties [19].

Calibrating dynamical models on experimental data time series is a central task in computational systems biology. When numerical values for model parameters can be found to fit the data, the model can be used to make predictions, whereas the absence of any good fit may suggest to revisit the structure of the model and gain new insights in the biology of the system. Temporal logic provides a formal framework to deal with imprecise data and specify a wide variety of dynamical behaviors. It can be used to extract information from numerical traces coming from either experimental data or model simulations, and to specify the expected behaviors for model calibration. The computation time of the different methods depends on the number of points in the trace so the question of trace simplification is important to improve their performance. In [19] we study this problem and provide a series of trace simplifications which are correct to perform for some common temporal logic formulae. We give some general soundness theorems, and apply this approach to period and phase constraints on the circadian clock and the cell cycle. In this application, temporal logic patterns are used to compute the relevant characteristics of the experimental traces, and to measure the adequacy of the model to its specification on simulation traces. Speed-ups by several orders of magnitude are obtained by trace simplification even when produced by smart numerical integration methods.

6.4. Highlight: Modeling Dynamics of Cell-to-Cell Variability in TRAIL-induced Apoptosis Explains Fractional Killing and Predicts Reversible Resistance

Participants: Grégory Batt, François Bertaux, Szymon Stoma.

Isogenic cells sensing identical external signals can take markedly different decisions. Such decisions often correlate with pre-existing cell-to-cell differences in protein levels. When not neglected in signal transduction models, these differences are accounted for in a static manner, by assuming randomly distributed initial protein levels. However, this approach ignores the *a priori* non-trivial interplay between signal transduction and the source of this cell-to-cell variability: temporal fluctuations of protein levels in individual cells, driven by noisy synthesis and degradation. Thus, modeling protein fluctuations, rather than their consequences on the initial population heterogeneity, would set the quantitative analysis of signal transduction on firmer grounds. Adopting this dynamical view on cell-to-cell differences amounts to recast extrinsic variability into intrinsic noise. In collaboration with Dirk Drasdo (EPI Mmaba), we proposed a generic approach to merge, in a systematic and principled manner, signal transduction models with stochastic protein turnover models. When applied to an established kinetic model of TRAIL-induced apoptosis, our approach markedly increased model prediction capabilities [4]. We obtained a mechanistic explanation of yet-unexplained observations on fractional killing and non-trivial robust predictions of the temporal evolution of cell resistance to TRAIL in HeLa cells. Our results provide an alternative explanation to survival via induction of survival pathways since no TRAIL-induced regulations are needed and suggest that short-lived anti-apoptotic protein Mcl1 exhibit large and rare fluctuations. More generally, our results highlight the importance of accounting for stochastic protein turnover to quantitatively understand signal transduction over extended durations, and imply that fluctuations of short-lived proteins deserve particular attention. [4]

6.5. Towards Real-time Control of Gene Expression at the Single Cell Level: A Stochastic Control Approach

Participants: Grégory Batt, Pascal Hersen.

Recent works have demonstrated the experimental feasibility of real-time gene expression control based on deterministic controllers. By taking control of the level of intracellular proteins, one can probe single-cell dynamics with unprecedented flexibility. However, single-cell dynamics are stochastic in nature, and a control framework explicitly accounting for this variability is presently lacking. In [21], we devised a stochastic control framework, based on Model Predictive Control, which fills this gap.

Based on a stochastic modelling of the gene response dynamics, our approach combined a full state-feedback receding-horizon controller with a real-time estimation method that compensated for unobserved state variables. Using previously developed models of osmostress-inducible gene expression in yeast, we showed *in silico* that our stochastic control approach outperformed deterministic control design in the regulation of single cells. This contribution lead to envision the application of the proposed framework to wet lab experiments in yeast.

This work was done in collaboration with Alfonso Carta (EPI BIOCORE), Eugenio Cinquemani (EPI IBIS), Lakshmeesh Maruthi and Ilya Tkachev (TU Delft), and Alessandro Abate (Oxford U).

6.6. A Platform for Rapid Prototyping of Synthetic Gene Networks in Mammalian Cells

Participants: Grégory Batt, Xavier Duportet, Pascal Hersen.

Mammalian synthetic biology may provide novel therapeutic strategies, help decipher new paths for drug discovery and facilitate synthesis of valuable molecules. Yet, our capacity to genetically program cells is currently hampered by the lack of efficient approaches to streamline the design, construction and screening of synthetic gene networks. To address this problem, we developed a framework for modular and combinatorial assembly of functional (multi)gene expression vectors and showed their efficient and specific targeted integration into a well-defined chromosomal context in mammalian cells.

In [7], in collaboration with the Weiss lab and the MSC lab, we demonstrated the potential of this framework by assembling and integrating different functional mammalian regulatory networks including the largest gene circuit built and chromosomally integrated to date (6 transcription units, 27kb), encoding an inducible memory device. Using a library of 18 different circuits as a proof of concept, we also demonstrated that our method enabled one-pot/single-flask chromosomal integration and screening of circuit libraries. This rapid and powerful prototyping platform is well suited for comparative studies of genetic regulatory elements, genes and multi-gene circuits as well as facile development of libraries of isogenic engineered cell lines.

6.7. Reconfigurable Circuitry in Biochemical Systems

Participants: Hui-Ju Chiang, François Fages, Sylvain Soliman.

Realizing complex systems within a biochemical environment is a common pursuit in synthetic biology. Such systems achieve certain computation through properly designed biochemical reactions. Despite fruitful progress being made, most existing reaction designs have fixed target functionality. Their lack of reconfigurability can be disadvantageous, especially when a system has to adapt to a varying biochemical environment.

When control systems are of concern, linear control is one of the most widely applied control methods. Any linear control system can be realized with three elementary building blocks: integration, gain, and summation. Realizing linear control with biochemical reactions has been proposed in previous work, where reaction rates of the underlying reactions play a key role to achieve the desired building blocks. Essentially the reaction rates have to be matched exactly, and it imposes serious practicality restriction because in reality the reaction rates of available reactions are predetermined and can be limited. In [16] we devise a mechanism to make linear control systems configurable by adding auxiliary species as control knobs. The concentrations of the auxiliary species can be adjusted not only to compensate reaction rate mismatch, but also to reconfigure different control systems out of the same control architecture.

Furthermore, in [15] we propose an analog approach to economically construct a reconfigurable logic circuit similar to a silicon based field programmable gate array (FPGA). The effective “logic” and “interconnect” of the circuit can be dynamically reconfigured by controlling the concentrations of certain knob species. We study a potential biomedical application of our reconfigurable circuitry to disease diagnosis and therapy at a molecular level.

6.8. Inferring Reaction Systems from Ordinary Differential Equations

Participants: François Fages, Steven Gay, Sylvain Soliman.

In Mathematical Biology, many dynamical models of biochemical reaction systems are presented with Ordinary Differential Equations (ODE). Once kinetic parameter values are fixed, this simple mathematical formalism completely defines the dynamical behavior of a system of biochemical reactions and provides powerful tools for deterministic simulations, parameter sensitivity analysis, bifurcation analysis, etc. However, without requiring any information on the reaction kinetics and parameter values, various qualitative analyses can be performed using the structure of the reactions, provided the reactants, products and modifiers of each reaction are precisely defined. In order to apply these structural methods to parametric ODE models, we study a mathematical condition for expressing the consistency between the structure and the kinetics of a reaction, without restricting to Mass Action law kinetics. This condition, satisfied in particular by standard kinetic laws, entails a remarkable property of independence of the influence graph from the kinetics of the reactions. We derive from this study a heuristic algorithm which, given a system of ODEs as input, computes a system of reactions with the same ODE semantics, by inferring well-formed reactions whenever possible. We show how this strategy is capable of automatically curating the writing of ODE models in SBML, and present some statistics obtained on the model repository biomodels.net [8].

6.9. Model Reductions by Tropical Equilibration

Participants: François Fages, Sylvain Soliman.

Model reduction is a central topic in systems biology and dynamical systems theory, for reducing the complexity of detailed models, finding important parameters, and developing multi-scale models for instance. While singular perturbation theory is a standard mathematical tool to analyze the different time scales of a dynamical system and decompose the system accordingly, tropical methods provide a simple algebraic framework to perform these analyses systematically in polynomial systems. The crux of these methods is in the computation of tropical equilibrations. In [11] we show that constraint-based methods, using reified constraints for expressing the equilibration conditions, make it possible to numerically solve non-linear tropical equilibration problems, out of reach of standard computation methods. We illustrate this approach first with the detailed reduction of a simple biochemical mechanism, the Michaelis-Menten enzymatic reaction model, and second, with large-scale performance figures obtained on the http://biomodels.net webcite repository.

6.10. Model Reductions by Subgraph Epimorphisms

Participants: François Fages, Steven Gay, Thierry Martinez, Sylvain Soliman.

In [9] we follow another route based on a purely structural method and study the problem of deciding the existence of a subgraph epimorphism between two graphs. Our interest in this variant of graph matching problem stems from the study of model reductions in systems biology, where large systems of biochemical reactions can be naturally represented by bipartite digraphs of species and reactions. In this setting, model reduction can be formalized as the existence of a sequence of vertex deletion and merge operations that transforms a first reaction graph into a second graph. This problem is in turn equivalent to the existence of a subgraph (corresponding to delete operations) epimorphism (i.e. surjective homomorphism, corresponding to merge operations) from the first graph to the second. In this paper, we study theoretical properties of subgraph epimorphisms in general directed graphs. We first characterize subgraph epimorphisms (SEPI), subgraph isomorphisms (SISO) and graph epimorphisms (EPI) in terms of graph transformation operations. Then we study the graph distance measures induced by these transformations. We show that they define metrics on graphs and compare them. On the algorithmic side, we show that the SEPI existence problem is NP-complete by reduction of SAT, and present a constraint satisfaction algorithm that has been successfully used to solve practical SEPI problems on a large benchmark of reaction graphs from systems biology.

6.11. Temporal Logic Modeling of Dynamical Behaviors: First-Order Patterns and Solvers

Participants: François Fages, Pauline Traynard, Sylvain Soliman.

We have written a book chapter [20] to describe how quantitative temporal logic formulae can be used to formalize imprecise dynamical behaviors of biological systems, and how such a formal specification of experimental observations can be used to calibrate models to real data, in a more versatile way than with curve fitting algorithms, and with more efficient dedicated solvers than with generic temporal logic solvers.

Based on this article, we investigated the correctness of various trace simplification methods, as mentionned in the highlight section above [19].

6.12. Logical Modeling of the Mammalian Cell Cycle

Participants: François Fages, Pauline Traynard, Denis Thieffry.

The molecular networks controlling cell cycle progression in various organisms have been previously modelled, predominantly using differential equations. However, this approach meets various difficulties as one tries to include additional regulatory components and mechanisms. This led to the development of qualitative dynamical models based on Boolean or multilevel frameworks, which are easier to define, simulate, analyse and compose. In a poster presented at ECCB 2014, we revisit the Boolean model of Fauré et al. for the core network controlling G/S transition in mammalian cell cycle, taking into account recent advances in the characterisation of the underlying molecular networks to obtain a better qualitative consistency between model simulations and documented mutants features. In particular, we introduced Skp2, the substrate recruiting component of the SCFSkp2 complex, which targets cell cycle control elements, such as p27, and is repressed by the tumour suppressor protein Rb. Furthermore, to supersede the limitations inherent to the Boolean simplifications, we have considered the association of multilevel logical components with key cell cycle regulators, including the tumour suppressor protein Rb. Indeed, it is well established that differently phosphorylated forms of Rb result in different effects on other components of the network, which can be faithfully modelled using a multilevel rather than a Boolean variable. To evaluate the dynamical properties of the resulting models, we perform synchronous and asynchronous simulations using the software GINsim (http://www.ginsim.org), for both the wild-type case and documented perturbations (e.g. combinations of loss- or gain-of-function mutations). In addition, we have designed a series of temporal logic queries (expressed in the CTL language), which enable an efficient and automatic verification of key dynamical properties (existence of a cyclic attractor or of a stable state, conditions on the order of changes of component levels, etc.), using the popular symbolic model checker NuSMV. This strategy greatly facilitates the dynamical analysis of increasingly detailed and complex cell cycle models. Our goal is to obtain a core cell cycle model consistent with the most relevant experimental results on mammalian cells, which will then be used as a module in more comprehensive cellular models, including cross-talks with the circadian clock network and key signalling pathways, whose deregulation underlies the development of various cancers.

6.13. A Greedy Heuristic for Optimizing Metro Regenerative Energy Usage compared to CMA-ES and MILP

Participants: François Fages, David Fournier.

When the regenerative braking energy cannot be stored by the metro producing it, it has to be used instantaneously on the network, otherwise it is lost. In this case, the accelerating and braking trains need be synchronized to fully benefit from the regenerative energy, and a metro timetable is energetically optimized when all the regenerative braking is utilized to power other trains. This synchronization consists in lining up each braking train with an accelerating one in its neighbourhood. Doing so, the latter will benefit from the regenerative energy of the former. In [17], [3] a fast greedy heuristic is proposed to tackle the problem of minimizing the energy consumption of a metro timetable by modifying solely the dwell times in stations. This heuristic is compared to a state-of-the-art meta heuristic called the covariance matrix adaptation evolution strategy (CMA-ES) and shows similar results with much faster computation time. Finally, it is shown that a run of the algorithm on a full timetable may reduce its energy consumption by 5.1%.

LINKMEDIA Project-Team

6. New Results

6.1. Highlights of the Year

BEST PAPER AWARD :

[21] **European Symposium on Research in Computer Security**. R. BALU, T. FURON, S. GAMBS.

6.2. Unsupervised motif discovery

6.2.1. Clustering by diverting supervised machine learning

Participants: Vincent Claveau, Patrick Gros, Abir Ncibi.

Knowledge discovery aims at bringing out coherent groups of objects and is usually based on clustering which necessitates defining a notion of similarity between objects. In practice, this strong prior is often neither possible nor welcome. We proposed to divert supervised machine learning (ML) techniques in order to calculate, indirectly and without supervision, similarities among objects. Our approach consists in generating artificial labeling problems on the data to reveal regularities between objects through their labeling. In [28], we show how this framework can be implemented and experimented on two information extraction/discovery tasks concerned with named entities. The ML technique diverted to exhibit similarities between with the named entities are the Conditional Random Fields. This same method can also be applied with less common ML techniques: In [59], we show that Inductive Logic Programming can also be used to cluster complex data. Thanks to the ability of ILP to handle data that cannot be expressed under the usual attribute-value representation, we use it to make emerge clusters of TV broadcasts based only on their broadcasting information (date, time, length, etc.).

6.2.2. Spoken term discovery applied to audio thumbnailing

Participants: Sébastien Campion, Guillaume Gravier.

We evaluated a system to create audio thumbnails of spoken content, i.e., short audio summaries representative of the entire content, without resorting to a lexical representation. As an alternative to searching for relevant words and phrases in a transcript, unsupervised motif discovery is used to find short, word-like, repeating fragments at the signal level without acoustic models. The output of the word discovery algorithm is exploited via a maximum motif coverage criterion to generate a thumbnail in an extractive manner. A limited number of relevant segments are chosen within the data so as to include the maximum number of motifs while remaining short enough and intelligible. Evaluation is performed on broadcast news reports with a panel of human listeners judging the quality of the thumbnails. Results indicate that motif-based thumbnails stand between random thumbnails and ASR-based keywords, however still far behind thumbnails and keywords humanly authored [34].

6.2.3. Unsupervised video structure mining with grammatical inference

Participants: Guillaume Gravier, Bingqing Qu.

In collaboration with Jean Carrive and Félicien Vallet, Institut National de l'Audiovisuel.

Unsupervised approaches were introduced a few years ago to analyze the structure of TV programs, relying on the discovery of repeated elements within a program or across multiple episodes of the same program. These methods can discover key repeating elements, such as jingles and separators, however they cannot infer the entire structure of a program. In [48], we studied a hierarchical use of grammatical inference to yield a temporal grammar of a program from a collection of episodes, discovering both the vocabulary of the grammar and the temporal organization of the words from the vocabulary. Using a set of basic event detectors and simple filtering techniques to detect repeating elements of interest, a symbolic representation of each episode is derived based on minimal domain knowledge. Grammatical inference based on multiple sequence alignment is then used in a hierarchical manner to provide a temporal grammar of the program at various levels of details.

6.2.4. Efficient indexing for content retrieval

Participants: Raghavendran Balu, Teddy Furon, Hervé Jégou.

In collaboration with Miajing Shi during, visiting Ph. D. student from Pekin University.

Many nearest neighbor search algorithms rely on encoding real vectors into binary vectors. The most common strategy projects the vectors onto random directions and takes the sign to produce so-called sketches. In [22], we discuss the sub-optimality of this choice, and propose a better encoding strategy based on the quantization and reconstruction points of view. Our second contribution is a novel asymmetric estimator for the cosine similarity. Similar to previous asymmetric schemes, the query is not quantized and the similarity is computed in the compressed domain. We tackled the same similarity estimation problem with a rather different approach in [52], where we assume that only a few vectors of the database, so-called heavy hitters, have a similarity to the query that significantly deviates from 0. For this purpose, we have introduced a group testing framework for detecting large similarities between high-dimensional vectors, such as descriptors used in state-of-the-art description of multimedia documents. We produce a set of group representations that jointly encode several vectors into a single one, in the spirit of group testing approaches. By comparing a query vector to several of these intermediate representations, we screen the large values taken by the similarities between the query and all the vectors, at a fraction of the cost of exhaustive similarity calculation. Unlike concurrent indexing methods that suffer from the curse of dimensionality, our method exploits the properties of high-dimensional spaces.

6.3. Multimedia content description and structuring

6.3.1. Linguistic knowledge extraction

6.3.1.1. Identifying events in texts

Participant: Vincent Claveau.

In collaboration with Béatrice Arnulphy, former team member now with ANR, Xavier Tannier and Anne Vilnat, LIMSI.

Identifying events from texts is an information extraction task necessary for many NLP applications. Through the TimeML specifications and TempEval challenges, it has received some attention in the last years, yet, no reference result was available for French. In [20], we tried to fill this gap by proposing several event extraction systems, combining for instance Conditional Random Fields, language modeling and k-nearest-neighbors. These systems are evaluated on French corpora and compared with state-of-the-art methods on English. The very good results obtained on both languages validate our whole approach and set new standard for French.

6.3.1.2. Morpho-semantic analysis of terms

Participants: Vincent Claveau, Ewa Kijak.

In most Indo-European languages, many biomedical terms are rich morphological structures composed of several constituents mainly originating from Greek or Latin. The interpretation of these compounds are keystones to access information. Following our work on morphology in the biomedical domain, we proposed different techniques to generate probabilistic morph-semantic resources and we show how these alignment information can be used for segmenting compounds, attaching semantic interpretation to each part, proposing definitions (gloses) of the compounds [26]. When possible, these tasks are compared with state-of-the-art tools, and the results show the interest of our automatically built probabilistic resources.

6.3.1.3. Distributional semantics

Participants: Vincent Claveau, Ewa Kijak.

In collaboration with Olivier Ferret, CEA-LIST.

We addressed the issue of building and improving a distributional thesaurus. We first show that existing tools from the information retrieval domain can be directly used in order to build a thesaurus with state-of-the-art performance. Secondly, we focus more specifically on improving the obtained thesaurus, seen as a graph of k-nearest neighbors. By exploiting information about the neighborhood contained in this graph, we propose several contributions. 1) We show how the lists of neighbors can be globally improved by examining the reciprocity of the neighboring relation, that is, the fact that a word can be close to another and vice-versa. 2) We also propose a method to associate a confidence score to any lists of nearest neighbors (i.e., any entry of the thesaurus). 3) Last, we demonstrate how these confidence scores can be used to reorder the closest neighbors of a word. These different contributions are validated through experiments and offer significant improvement over the state-of-the-art [27], [60].

6.3.2. OCR and speech content-based description

6.3.2.1. Use of stress information for robust speech recognition

Participant: Guillaume Gravier.

In collaboration with S. Ziegler, PANAMA Inria team and Laboratoire de Sciences Cognitives et Psycholinguistique.

[44] presents a study on the robustness of stress information for automatic speech recognition in the presence of noise. The syllable stress, extracted from the speech signal, was integrated in the recognition process by means of a previously proposed decoding method. Experiments were conducted for several signal-to-noise ratio conditions and the results show that stress information is robust in the presence of medium to low noise. This was found to be true both when syllable boundary information was used for stress detection and when this information was not available. Furthermore, the obtained relative improvement increased with a decrease in signal quality, indicating that the stressed parts of the signal can be considered islands of reliability.

6.3.2.2. Boosting bonsai trees for handwritten/printed text discrimination

Participant: Christian Raymond.

In collaboration with Yann Ricquebourg, Baptiste Poirriez, Aurélie Lemaitre and Bertrand Coüasnon, IRISA.

Boosting over decision-stumps proved its efficiency in natural language processing, essentially with symbolic features, and its good properties (fast, few and not critical parameters, not sensitive to overfitting) could be of great interest in the numeric world of pixel images. In [51], we investigated the use of boosting over small decision trees in image classification processing for the discrimination of handwritten/printed text. We conducted experiments to compare with usual SVM-based classification revealing convincing results with very close performance, but with faster predictions and behaving far less as a black-box. Those promising results tend to make use of this classifier in more complex recognition tasks like multiclass problems.

6.3.2.3. Speaker role detection from spoken document

Participant: Christian Raymond.

In collaboration with LIMSI and LIUM.

In [40] and [41], we tackle the problem of speaker role detection in broadcast news shows. In the literature, many proposed solutions are based on the combination of various features coming from acoustic, lexical and semantic information with a machine learning algorithm. Many previous studies mention the use of boosting over decision stumps to combine efficiently these features. We proposed a modification of this state-of-the-art machine learning algorithm changing the weak learner (decision stumps) by small decision trees, denoted bonsai trees. Experiments show that using bonsai trees as weak learners for the boosting algorithm largely reduces both system error rate and learning time.

6.3.3. Image and video description and classification

6.3.3.1. Fine-grain image classification

Participants: Teddy Furon, Philippe-Henri Gosselin, Hervé Jégou.

In collaboration with Xerox Research Center Europe.

We have addressed the problem of instance classification: our goal is to annotate images with tags corresponding to objects classes which exhibit small intra-class variations such as logos, products or landmarks. Our first contribution on image classification [13] describes the processing pipeline, which has won FGCOMP challenge associated with Imagenet. It improves a standard method based on Fisher vectors to adapt it to the context of fine-grained classes, where the difference between classes rely on few but typical visual differences. On the same task, we have proposed a novel algorithm [39] for the selection of class-specific *prototypes* which are used in a voting-based classification scheme.

6.3.3.2. Aggregation of local descriptors

Participants: Teddy Furon, Hervé Jégou, Giorgos Tolias.

In collaboration with the University of Oxford.

For unsupervised particular object and image recognition, we have considered the design of a single vector representation for an image that embeds and aggregates a set of local patch descriptors such as SIFT. In [36], we make two contributions, both aimed at regularizing the individual contributions of the local descriptors in the final representation. The first is a novel embedding method that avoids the dependency on absolute distances by encoding directions. The second contribution is a “democratization” strategy that further limits the interaction of unrelated descriptors in the aggregation stage. In [36], we addressed another issue inherent to existing encoding algorithms: Image search systems based on local descriptors typically achieve orientation invariance by aligning the patches on their dominant orientations. This choice introduces too much invariance because it does not guarantee that the patches are rotated consistently. To address this problem, we have introduced another aggregation strategy of local descriptors that achieves this covariance property by jointly encoding the angle in the aggregation stage in a continuous manner. It is combined with an efficient monomial embedding to provide a codebook-free method to aggregate local descriptors into a single vector representation.

6.3.3.3. Action localization in videos

Participants: Mihir Jain, Hervé Jégou.

In collaboration with the University of Amsterdam and the project-team SERPICO.

We have tackled the problem of action localization in videos [35], where the objective is to determine when and where certain actions appear. We introduce a sampling strategy, called tubelets and inspired a method recently introduced for image detection. It drastically reduces the number of hypotheses that are likely to include the action of interest. By using super-voxels and employing a criterion that reflects how action related motion deviates from background motion, the method is specifically adapted to 2D+t sequences and establishes the new state-of-the-art for action localization on the public datasets UCF Sports and MSR-II.

6.3.4. Text description for information retrieval

Participants: Vincent Claveau, Sébastien Le Maguer.

In collaboration with Natalia Grabar, STL UMR8163, and Thierry Hamon, LIMSI

Following previous work, we investigated the interest of “bag of bags of features” representation for texts in an vector-space information retrieval setting. Each text is thus represented as a bag of vector. With this representation, computing the similarity between two texts necessitates to aggregate every vector to vector similarity for the two bags. In [58], we examine the expected properties of such an aggregation function and show their influence through different experiments. When some specific conditions are met, we show that the gains over standard representation can be very important.

With a team composed with members of TEXMEX/LINKMEDIA, LIMSI and STL, we have participated to the biomedical information retrieval challenge proposed in the framework of CLEF eHealth [25]. For this first participation, our approach relies on a state-of-the-art IR system called Indri, based on statistical language modeling, and on semantic resources. The purpose of semantic resources and methods is to manage the term variation such as synonyms, morpho-syntactic variants, abbreviation or nested terms. Different combinations of resources and Indri settings are explored, mostly based on query expansion. We obtained good overall results (3rd in terms of MAP) and confirmed the interest of query expansion to retrieve a maximum of relevant documents.

6.4. Linking, navigation and analytics

6.4.1. NLP-driven hyperlink construction in broadcast videos

Participants: Rémi Bois, Vincent Claveau, Guillaume Gravier, Pascale Sébillot, Anca-Roxana Şimon.

In collaboration with Sien Moens, Katholieke Universiteit Leuven, Éric Jamet and Martin Ragot, Univ. Rennes 2.

The hyperlinking sub-task of the MediaEval Search and Hyperlinking task aims at creating hyperlinks between predefined anchor segments, i.e., fragments of videos, and short related video segments, called targets, that have to be automatically extracted from videos of a large collection. Capitalizing on the experience acquired in previous participations [54], we proposed a two step approach exploiting speech material: Potential target segments are first generated relying on a topic segmentation technique; For each anchor, the best 20 target segments are then selected according to two distinct strategies. The first strategy focuses on the identification of very similar targets using n-grams and named entities, while the second one makes use of an intermediate structure built from topic models, which offers the possibility to control serendipity and to explain the links created [53].

In 2014, we also initiated the CominLabs project "Linking media in acceptable hypergraphs" dedicated to the creation of explicit and meaningful links between multimedia documents or fragments of documents. Two main issues were adressed: The construction of a corpus, composed of audio and video news, reports and debates, newspapers and blog websites, as well as social networks; A preliminary study of the perceived usefulness of various types of links by end-users.

6.4.2. Analytics in collections of art critics

Participant: Vincent Claveau.

In collaboration with Fabienne Moreau and Nicolas Thély, Univ. Rennes 2.

We aim at exploiting text mining techniques in the service of digital humanities, and more precisely in the field of art criticism. It relies on a collaboration between our team, linguists and art and aesthetics specialists. In preliminary work [56], we adapted term extraction, named entity recognition and information retrieval techniques to this field to extract multiple linguistic clues from art review articles. Future work will make the most of these clues and clustering approaches to build a navigable and structured collection of the articles.

6.4.3. Data models for navigation

Participant: Laurent Amsaleg.

In collaboration with Björn Þór Jónsson, Grímur Tómasson, Hlynur Sigurþórsson, Áslaug Eríksdóttir and Marta Kristin Larusdottir, School of Computer Science, Reykjavík University.

Digital photo collections—personal, professional, or social—have been growing ever larger, leaving users overwhelmed. It is therefore increasingly important to provide effective browsing tools for photo collections. Learning from the resounding success of multi-dimensional analysis (MDA) in the business intelligence community for On-Line Analytical Processing (OLAP) applications, we propose a multi-dimensional model for media browsing, called M^3, that combines MDA concepts with concepts from faceted browsing. We present the data model and describe preliminary evaluations, made using server and client prototypes, which indicate that users find the model useful and easy to use [38]. A photo navigation prototype was demonstrated at the Intl. Conf. on Multimedia Modeling [37].

6.4.4. Exploiting k-nn graphs for image retrieval

Participants: Laurent Amsaleg, Hervé Jégou, Giorgos Tolias.

We have proposed two techniques exploiting the relationship between the images with a collection. In [29], we revisit how to exploit the k-reciprocal nearest neighbors to produce, for a given query, a neighborhood that improves over the one obtained with the original metric. This strategy is simpler than concurrent prior work, yet it is both effective and less sensitive to parameters. Second, we propose to employ measures defined on sets of shared nearest neighbors in order to re-rank the shortlist. Both methods are simple, yet they significantly improve the accuracy of image search engines on standard benchmarks. We also introduced a query expansion technique [18] for image search that is faster and more precise than the existing ones. The expansion generates an enriched representation which refines the initial local descriptors *individually* by aggregating those of the database, while new descriptors are produced from the images that are deemed relevant. The technique has two computational advantages over other query expansion techniques. First, the size of the enriched representation is comparable to that of the initial query. Second, the technique is effective even without using any geometry, in which case searching a database comprising 105k images typically takes 80 ms on a desktop machine. Overall, our technique significantly outperforms the visual query expansion state of the art on popular benchmarks.

6.5. Content-based information retrieval and security

6.5.1. Analysis of privacy preserving data aggregation for recommendation systems

Participants: Raghavendran Balu, Teddy Furon.

Work under the Alcaltel-Lucent / Inria common Lab, in collaboration with Armen Aghasaryan, Dimitre Davidov, Makram Bouzid (ALU) and Sébastien Gambs (Inria Rennes Cidre team project).

We consider personalized recommendation systems in which before publication, the profile of a user is sanitized by a non-interactive mechanism compliant with the concept of differential privacy. We analyze two existing schemes offering a differentially private representation of profiles: BLIP (BLoom-and-flIP) and JLT (Johnson-Lindenstrauss Transform). For assessing their security levels, we play the role of an adversary aiming at reconstructing a user profile [21]. We compare two inference attacks, namely single and joint decoding. The first one decides of the presence of a single item in the profile, and sequentially explores all the item set. The second one decides whether a subset of items is likely to be the user profile, and considers all the possible subsets. Our contributions are a theoretical analysis as well as a practical implementation of both attacks, which were evaluated on datasets of real user profiles. The results obtained clearly demonstrates that joint decoding is the most powerful attack, while also giving useful insights on how to set the differential privacy parameter ϵ.

6.5.2. Content based image retrieval with privacy

Participants: Laurent Amsaleg, Teddy Furon, Li Weng.

Work initiated during a collaboration with A. Morton, L. Weng (with LINKMEDIA *since May 2014) and S. Marchand-Maillet, Université de Genève.*

We propose a privacy protection framework for large-scale content-based information retrieval. It offers two layers of protection. First, robust hash values are used as queries to prevent revealing original content or features. Second, the client can choose to omit certain bits in a hash value to further increase the ambiguity for the server. Due to the reduced information, it is computationally difficult for the server to know the client's interest. The server has to return the hash values of all possible candidates to the client. The client performs a search within the candidate list to find the best match. Since only hash values are exchanged between the client and the server, the privacy of both parties is protected. We introduce the concept of tunable privacy, where the privacy protection level can be adjusted according to a policy. It is realized through hash-based piece-wise inverted indexing. The idea is to divide a feature vector into pieces and index each piece with a sub-hash value. Each sub-hash value is associated with an inverted index list. The framework has been extensively tested using a large image database. We have evaluated both retrieval performance and privacy-preserving performance for a particular content identification application.

6.5.3. Privacy protection in face recognition

Participant: Teddy Furon.

In collaboration with B. Bhattarai, A. Mignon, F. Jurie, GREYC, Université of Caen.

We investigated a new approach for de-identifying face images, *i.e.* for preventing automatic matching with public face collections. The overall motivation is to offer tools for privacy protection on social networks. We address this question by drawing a parallel between face de-identification and oracle attacks in digital watermarking. In our case, the identity of the face is seen as the watermark to be removed. Inspired by oracle attacks, we forge de-identified faces by superimposing a collection of carefully designed noise patterns onto the original face. The modification of the image is controlled to minimize the probability of good recognition while minimizing the distortion. In addition, these de-identified images are by construction made robust to counter attacks such as blurring. We present an experimental validation in which we de-identify LFW faces and show that resulting images are still recognized by human beings while deceiving a state-of-the-art face recognition algorithm [23].

6.5.4. Tardos code in practice

Participant: Teddy Furon.

Joint work with the Inria supported start-up LAMARK.

We deal with active fingerprinting a.k.a. traitor tracing where a collusion of dishonest users merges their individual versions of a content to yield a pirated copy. The Tardos codes are one of the most powerful tools to fight against such collusion process by identifying the colluders. Instead of studying as usual the necessary and sufficient code length in a theoretical setup, we adopt the point of view of the practitioner. We call this the *operational mode*, *i.e.* a practical setup where a Tardos code has already been deployed and a pirated copy has been found. This new paradigm shows that the known bounds on the probability of accusing an innocent in the theoretical setup are way too pessimistic. Indeed the practitioner can resort to much tighter bounds because the problem is fundamentally much simpler under the operational mode. In the end, we benchmark under the operational mode several single decoders recently proposed in the literature [32].

LINKS Team

5. New Results

5.1. Highlights of the Year

In the objective Querying Heterogeneous Linked Data, Slawomir Staworko and Iovka Boneva have developed new ways to define schema for Graph Database and RDF [19]. This work has been influencing a group work of W3C on defining a schema for the DF format. This work is a continuation of [3] (by Iovka Boneva, Radu Ciucanu and Slawomir Staworko) developping a new schema for unordered trees over XML. Due to these works, Boneva is now a member of the Data Shapes Working Group which mission is to produce a language for defining structural constraints on RDF graphs. http://www.w3.org/2014/data-shapes/charter

In the objective Managing Dynamic Linked Data, the main break through is the development of QuixPath that now covers 100 per cent of the XPathMark, a W3C benchmark for the language Xpath (querying XML trees). In particular, it includes aggregation operators, joins and arithmetics operations. The core of QuixPath is based on techniques presented in [6] (by Tom Sebastian, Denis Debardieux and Joachim Niehren).

In the objective Linking Data Graphs, different methods have been developped to learn queries over graph. More precisely, the queries learned are conjunctive queries with joins. These techniques have been presented in [13] and demonstrated in [4] at the conference VLDB.

5.2. Querying Heterogeneous Linked Data

Angela Bonifati, Gianvito Summa, Esther Pacitt (U Montpellier 2)i and Fady Draidi (U Montpellier 2) [5] consider peer-to-peer data management systems (PDMS), where each peer maintains mappings between its schema and some acquaintances, along with social links with peer friends. In this context, the goal is reformulating conjunctive queries from a peer's schema into other peer's schemas. Precisely, queries against a peer node are rewritten into queries against other nodes using schema mappings thus obtaining query rewritings. They propose a new notion of 'relevance' of a query with respect to a mapping that encompasses both a local relevance (the relevance of the query w.r.t. the mapping) and a global relevance (the relevance of the query w.r.t. the entire network). Based on this notion, they conceived a new query reformulation approach for social PDMS which achieves great accuracy and flexibility. This has been implemented and experimented in a prototype.

Pierre Bourhis, Andreas Morak and Andreas Pieris [14] investigated classes of queries for which the problem of open query answering of disjunctive guarded TGDs a decent complexity (e, g exptime). The complete picture of the complexity of answering (unions of) conjunctive queries under the main guarded-based classes of disjunctive existential rules has been recently settled. It has been shown that the problem is very hard, namely 2ExpTime-complete, even for fixed sets of rules expressed in lightweight formalisms.The central objective of the present paper is to understand whether simpler query languages (bounded tree width and acyclic queries) have a positive impact on the complexity of query answering under the main guarded-based classes of disjunctive existential rules.

In [3], a new formalism for schema for unordered trees have been developed. It is based on a notion of regular expressions of multisets of labels. Different problems of static analysis like emptiness and containment are studied and their complexity. Different simpler schema are studied leading to interesting complexity for the different studied problems. Finally, they study the expressive power of the proposed schema languages and compare them with yardstick languages of unordered trees (FO, MSO, and Presburger constraints) and DTDs under commutative closure. The results show that the proposed schema languages are capable of expressing many practical languages of unordered trees and enjoy desirable computational properties.

In [7], Adrian Boiret, Vincent Hugot and Joachim Niehren and Ralf Treinen (University Paris 7) proposes a notion deterministic tree automata for unordered trees. While the existing notions are well-investigated concerning expressiveness, they all lack a proper notion of determinism, which makes it difficult to distinguish subclasses of automata for which problems such as inclusion, equivalence, and minimization can be solved efficiently. In this paper, the authors propose and investigate different notions of ”horizontal determinism”, starting from automata for unranked trees in which the horizontal evaluation is performed by finite state automata.

5.3. Managing Dynamic Linked Data

Tom Sebastian, Denis Debarbieux, Olivier Gauwin (U Bordeaux), Joachim Niehren, Mohamed Zergaoui (Innovimax) [6] present new techniques to evaluate XPath queries on trees received in a streaming way. It introduce early nested word automata in order to approximate earliest query answering algorithms for nested word automata. The notion early query answering algorithm is based on stack-and-state sharing for running early nested word automata on all answer candidates with on-the-fly determinization. These techniques allow to implement a more important part of Xpath and outcome all the previous tools in coverage of XpathMark benchmark.

5.4. Linking Data Graphs

Angela Bonifati, Radu Ciucanu, Slawomir Staworko developed techniques to learn conjunctive queries from example given by a user. The main part is to infer joins between relations from the positive and negative tuples. Different techniques to deduce informative examples are presented and interestingly they can be done in polynomial time. The techniques are published in [13] and demonstrated in [4].

Grégoire Laurence, Aurélien Lemay, Joachim Niehren, Slawek Staworko, Marc Tommasi [16] explain how to learn sequential top-down tree-to- word transducers (STWs). First, they present a Myhill-Nerode characterization of the corresponding class of sequential tree-to-word transformations (STW). Next, they investigate what learning of stws means, identify fundamental obstacles, and propose a learning model with abstain. Finally, they present a polynomial learning algorithm

M3DISIM Team

6. New Results

6.1. Highlights of the Year

- Radomir Chabiniok recruited in starting research position (start Febr 2015);
- PhD Defense of Annabelle Collin;
- “Usine Nouvelle” article.

6.2. Modeling

6.2.1. Mechanics of collective unfolding

Participants: Matthieu Caruel [correspondant], Jean-Marc Allain [LMS], Lev Truskinovsky [LMS].

Mechanically induced unfolding of passive crosslinkers is a fundamental biological phenomenon encountered across the scales from individual macro-molecules to cytoskeletal actin networks. In this work we study a conceptual model of athermal load-induced unfolding and use a minimalistic setting allowing one to emphasize the role of long-range interactions while maintaining full analytical transparency. Our model can be viewed as a description of a parallel bundle of N bistable units confined between two shared rigid backbones that are loaded through a series spring. We show that the ground states in this model correspond to synchronized, single phase configurations where all individual units are either folded or unfolded. We then study the fine structure of the wiggly energy landscape along the reaction coordinate linking the two coherent states and describing the optimal mechanism of cooperative unfolding. Quite remarkably, our study shows the fundamental difference in the size and structure of the folding-unfolding energy barriers in the hard (fixed displacements) and soft (fixed forces) loading devices which persists in the continuum limit. We argue that both, the synchronization and the non-equivalence of the mechanical responses in hard and soft devices, have their origin in the dominance of long-range interactions. We then apply our minimal model to skeletal muscles where the power-stroke in acto-myosin crossbridges can be interpreted as passive folding. A quantitative analysis of the muscle model shows that the relative rigidity of myosin backbone provides the long-range interaction mechanism allowing the system to effectively synchronize the power-stroke in individual crossbridges even in the presence of thermal fluctuations. In view of the prototypical nature of the proposed model, our general conclusions pertain to a variety of other biological systems where elastic interactions are mediated by effective backbones.

6.2.2. Thermodynamical framework for modeling chemical-mechanical coupling in muscle contraction - Formulation and validation

Participants: Matthieu Caruel, Dominique Chapelle [correspondant], Philippe Moireau.

Muscle contraction occurs at the nanoscale of a hierarchical multi-scale structure with the attachment of so-called cross-bridges within sarcomeres, namely, the creation of chemical bonds between myosin heads and specific sites on actin filaments. A cross-bridge in itself can be seen as a special chemical entity having internal mechanical variables - or degrees of freedom - pertaining to the actual geometric configuration, which implies that the free energy of the cross-bridge - whether in an attached or unattached state - must be made dependent on these internal variables (T.L. Hill, Free Energy Transduction And Biochemical Cycle Kinetics, Dover, 2004). This provides a thermodynamical basis for modeling the complex interplay of chemical and mechanical phenomena at the sarcomere level. Within this framework we propose a muscle model with two mechanical variables associated with a cross-bridge. For the action of individual cross-bridges occurring at the nanometer scale, the energy provided by the Langevin thermostat cannot be neglected, and we therefore propose to endow the internal mechanical variables with stochastic dynamics. Important motivations for this modeling choice include the ability to represent (i) the so-called power-stroke phenomenon and (ii) short-time responses of a muscle, e.g. to load steps. Our approach allows for systematic treatment of the model energetics, and in particular one goal of the proposed description is to investigate the potential benefit in mechanical efficiency with systems including - in addition to chemically-induced transformations - thermally-induced conformational changes such as the power-stroke.

6.2.3. Mechanical modeling and numerical methods for poromechanics: Applications to cardiac perfusion

Participants: Bruno Burtschell, Dominique Chapelle [correspondant], Philippe Moireau.

We have previously formulated a rather general modeling framework of poromechanics – formulations that combine solid and fluid components to represent the behavior of a porous medium – to take into account large deformations and rapid fluid flows, see [16]. This allows to consider, in particular, the application of blood perfusion within the cardiac tissue, which features these specific complex phenomena, out of the scope of classical poromechanical models. One of our major objectives now, within the PhD of Bruno Burtschell, is to propose and assess some associated relevant numerical schemes.

Some existing algorithms of fluid-structure interaction, with which our poromechanics formulations feature deep similarities, have been implemented – in FreeFEM++, both in axisymmetric configuration and in 3D – and compared. Their numerical and theoretical analysis – consistency, convergence – has been performed. Then, the adaptation of these algorithms to our poromechanics formulations enabled us to propose a time discretisation well-fitted to our framework, and to present its energy stability analysis. Further perspectives include implementation and numerical validation of this scheme, including special care regarding space discretisation, then integration into FELISCE ("HappyHeart" module).

6.2.4. Personalized modeling for cardiac amyloidosis diagnosis

Participants: Alessandro Felder, Dominique Chapelle, Philippe Moireau, Jean-François Deux [Hôpital Henri Mondor], Thibault Damy [Hôpital Henri Mondor].

Cardiac amyloidosis is a condition induced by pathological deposition of amyloid proteins within muscle tissue and nerves, thus severely impairing the cardiac function and often requiring cardiac transplantation as the only available treatment. Our objective here in a first stage is to use our previously developed patient-specific modeling methodologies to analyse some clinical cases – based on actual patient data – to better apprehend the impact of the pathology on biomechanical properties. Further perspectives include the modeling of the protein deposition and associated tissue remodeling in order to predict the disease evolution in a patient-specific context. This work is performed in collaboration with medical doctors from Hôpital Henri Mondor (Créteil).

6.3. Numerical Analysis

6.3.1. Fourth-order energy-preserving locally implicit discretization for linear wave equations

Participants: Juliette Chabassier [Magique-3d], Sébastien Imperiale [correspondant].

A family of fourth-order coupled implicit-explicit time schemes has been developed. The spatial coupling is done at the boundaries of several non conforming meshes of regions in which we want to simulate propagating waves. A global discrete energy is shown to be preserved and leads to global fourth-order consistency. Numerical results in 1D and 2D have been produced to illustrate the good behavior of the schemes and their potential for the simulation of realistic highly heterogeneous media and strongly refined geometries, for which using an explicit scheme everywhere can be extremely penalizing. Accuracy up to fourth order reduces the numerical dispersion inherent to implicit methods used with a large time step, and makes this family of schemes attractive compared to second order accurate methods in time.

6.4. Model-Data Interaction

6.4.1. A Luenberger observer for reaction-diffusion models with front position data

Participants: Dominique Chapelle, Annabelle Collin, Philipe Moireau [correspondant].

We propose a Luenberger observer for reaction-diffusion models with propagating front features, and for data associated with the location of the front over time. Such models are considered in various application fields, such as electrophysiology, wild-land fire propagation and tumor growth modeling. Drawing our inspiration from image processing methods, we start by proposing an observer for the eikonal-curvature equation that can be derived from the reaction-diffusion model by an asymptotic expansion. We then carry over this observer to the underlying reaction-diffusion equation by an "inverse asymptotic analysis", and we show that the associated correction in the dynamics has a stabilizing effect for the linearized estimation error. We also discuss the extension to joint state-parameter estimation by using the earlier-proposed ROUKF strategy. We then illustrate and assess our proposed observer method with test problems pertaining to electrophysiology modeling, including with a realistic model of cardiac atria. Our numerical trials show that state estimation is directly very effective with the proposed Luenberger observer, while specific strategies are needed to accurately perform parameter estimation – as is usual with Kalman filtering used in a nonlinear setting – and we demonstrate two such successful strategies.

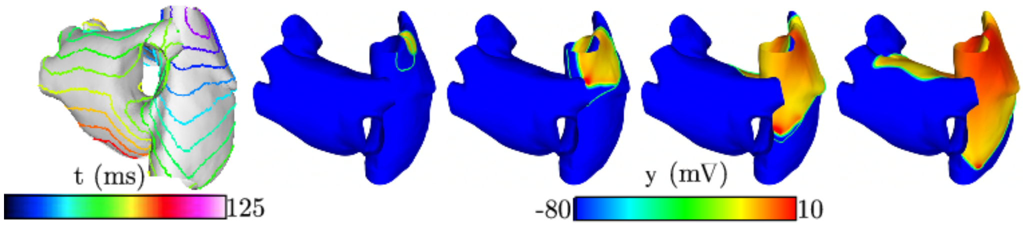


Figure 1. Collocated front data on an atria (left), and observer of the atrial electric activation pursuing the green front from a wrong initial condition (right, 4 time-steps)

6.4.2. Identification of weakly coupled multiphysics problems. Application to the inverse problem of electrocardiography

Participants: Cesare Corrado [Reo team], Jean-Frédéric Gerbeau [Reo team], Philippe Moireau [correspondant].

This work addresses the inverse problem of electrocardiography from a new perspective, by combining electrical and mechanical measurements. Our strategy relies on the definition of a model of the electromechanical contraction which is registered on ECG data but also on measured mechanical displacements of the heart tissue typically extracted from medical images. In this respect, we establish in this work the convergence of a sequential estimator which combines for such coupled problems various state of the art sequential data assimilation methods in a unified consistent and efficient framework. Indeed, we aggregate a Luenberger observer for the mechanical state and a Reduced-Order Unscented Kalman Filter applied on the parameters to be identified and a POD projection of the electrical state. Then using synthetic data we show the benefits of our approach for the estimation of the electrical state of the ventricles along the heart beat compared with more classical strategies which only consider an electrophysiological model with ECG measurements. Our numerical results actually show that the mechanical measurements improve the identifiability of the electrical problem allowing to reconstruct the electrical state of the coupled system more precisely. Therefore, this work is intended to be a first proof of concept, with theoretical justifications and numerical investigations, of the advantage of using available multi-modal observations for the estimation and identification of an electromechanical model of the heart.

6.4.3. Data assimilation for hyperbolic conservation laws. A Luenberger observer approach based on a kinetic description

Participants: Anne-Céline Boulanger [Ange team], Benoît Perthame [Mamba team], Philippe Moireau [correspondant], Jacques Sainte-Marie [Ange team].

Developing robust data assimilation methods for hyperbolic conservation laws is a challenging subject. Those PDEs indeed show no dissipation effects and the input of additional information in the model equations may introduce errors that propagate and create shocks. We propose a new approach based on the kinetic description of the conservation law. A kinetic equation is a first order partial differential equation in which the advection velocity is a free variable. In certain cases, it is possible to prove that the nonlinear conservation law is equivalent to a linear kinetic equation. Hence, data assimilation is carried out at the kinetic level, using a Luenberger observer also known as the nudging strategy in data assimilation. Assimilation then amounts to the handling of a BGK type equation. The advantage of this framework is that we deal with a single "linear" equation instead of a nonlinear system and it is easy to recover the macroscopic variables. The study is divided into several steps and essentially based on functional analysis techniques. First we prove the convergence of the model towards the data in case of complete observations in space and time. Second, we analyze the case of partial and noisy observations. To conclude, we validate our method with numerical results on Burgers equation and emphasize the advantages of this method with the more complex Saint-Venant system.

6.4.4. Optimal observer for parabolic problems

Participants: Karine Mauffrey, Philippe Moireau [correspondant].

We aim at proposing optimal observers strategies for reconstructing the solution of general systems of PDEs using available observations, including both wave-type equations and heat-like equations. The main objective of this work is to present a self-contained analysis. For a general parabolic system, we have established the exponential stability of the operator occurring in the equation satisfied by the error between the target and the optimal observer. The proof relies on two major hypotheses: an observability inequality satisfied by the observation operator and a controllability property for the modeling error operator by which model noises enter the dynamics (controllability property which is related to the invertibility of the solution of the associated infinite dimensional Riccati equation). Then we have tackled the discretisation questions and demonstrated that the discrete-time Kalman filter is an adequate discretization of the continuous-time Kalman filter. Finally we have also studied the strong formulation of the Kalman observer using a kernel representation of the Riccati operator.

6.4.5. Elastography by magnetic resonance imaging

Participants: Guillaume Bal [Columbia Unviersity], Cedric Bellis [LMA Marseille], Sébastien Imperiale [correspondant], Francois Monard [University of Washington- Seattle].

We have studied the potential application of elastography by Magnetic Resonance Imaging (MRI) within the framework of linear elasticity. We assume given internal full-field MRI measurements of the deformations of a non-homogeneous isotropic solid, and the aim is the quantitative reconstruction of the associated physical parameters. Upon using polluted measurements, a variational formulation is constructed, its inversion enabling the recovery of the parameters. The analysis of this inversion procedure provides existence and uniqueness results while the reconstruction stability with respect to the measurements is investigated. As the inversion procedure requires differentiating the measurements twice, a numerical differentiation scheme has been proposed and analyzed. It is based on a regularization that allows an optimally stable reconstruction of the sought parameters.

MADYNES Project-Team

6. New Results

6.1. Highlights of the Year

The following points of 2014 deserves to be highlighted:

- One new permanent member joined the MADYNES team: Jérôme François as Inria researcher.
- An IBM Faculty Award has been received by a team member (Rémi Badonnel, TELECOM Nancy) for his work on security and cloud computing.

BEST PAPER AWARD :
[21] **8th IFIP WG 6.6 International Conference on Autonomous Infrastructure, Management, and Security, AIMS 2014**. A. MAYZAUD, A. SEHGAL, R. BADONNEL, I. CHRISMENT, J. SCHÖNWÄLDER.

6.2. Monitoring

6.2.1. P2P network monitoring

Participants: Thibault Cholez [contact], Isabelle Chrisment, Olivier Festor.

Finishing a work started several years ago with our colleagues from the team Complex Network [0] at the LIP6, we published a final result on the comparison of paedophile activity in different P2P systems [5]. We designed a methodology for comparing KAD and eDonkey, two P2P systems among the most prominent ones and with different anonymity levels. We have detected paedophile-related queries with a previously validated tool and we proposed, for the first time, a large-scale comparison of paedophile activity in two different P2P systems.

We are also glad to have contributed to a book chapter in french on the uses and misuses of digital identities on the Internet [33]. It summarizes several years of work of the team, fighting against the Sybil attack in P2P networks in order to improve their security and quality of service.

6.2.2. Anonymous networks monitoring

Participants: Juan Pablo Timpanaro, Isabelle Chrisment [contact], Olivier Festor.

Anonymous networks have emerged to protect the privacy of network users. Large scale monitoring on these systems allows us to understand how they behave and which type of data is shared among users.

In 2014, we continued our research about the I2P anonymous network [0]. This network is optimized for anonymous web hosting and anonymous file-sharing. I2P's file-sharing community is highly active with users deploying their file-sharing applications on top of the network. I2P uses a variation of Onion routing, thus assuring the unlinkability between a user and its file-sharing application. In [26] we took the first step towards the linkability of users and applications in the I2P network. We conducted a group-based characterization, where we determine to what extent a group of users is responsible for the overall I2P's file-sharing activity. We used Pearson's coefficient to correlate users from two cities and the most used anonymous file-sharing application.

6.2.3. Smartphone usage monitoring

Participants: Vassili Rivron [contact], Mohammad Irfan Khan, Simon Charneau [Inria], Isabelle Chrisment.

Over the last few years the number of smartphone applications has increased enormously. In 2014, we passively collected smartphones usage logs in the wild by inviting the crowd to participate in the PRACTIC [0] contest and install our crowdsensing application to contribute anonymous smartphone usage logs, voluntarily and in the most natural settings (their own phone, own pricing plan) .

[0] http://www.complexnetworks.fr/
[0] http://i2p2.de
[0] http://beta.apisense.fr/practic

Complementary to sensing we also collected contextual information (social, demographic, professional) and information about users'perception via survey questionnaires built in the application or on the web.

This experiment used a crowd sensing platform called APISENSE ®[0] and developed by the Inria Spirals Team. It was carried out in the context of building a country-wide Internet observation platform in France, called Metroscope [0].

6.3. Security

6.3.1. Security Automation

Participants: Rémi Badonnel [contact], Martin Barrere, Gaëtan Hurel, Abdelkader Lahmadi, Olivier Festor.

The main research challenge addressed in this work is focused on enabling configuration security automation in dynamic networks and services.

A first part of our work in the year 2014 was centered on a strategy for remediating known vulnerabilities, formalizing the correction decision problem as a satisfiability or SAT problem [10]. From a proactive perspective, it should be able to decide which potential states could be dangerous. By specifying our vulnerability knowledge source (OVAL repository) as a propositional logical formula, we have fixed system properties that we cannot change and free those variables for which changes are available. We have introduced the X2CCDF language, built on top of XCCDF and OVAL, that allows us to express the impact of these changes over target systems. These descriptions can be used for analyzing the security impact of changes without actually changing the system. When this information is not available, we have considered the NETCONF protocol and its notion of candidate state where changes can be applied, analyzed and rolled back if necessary.

A second part of our work has been dedicated to the orchestration of security functions in the context of mobile smart environments [19]. Most of current security approaches for these environments are provided in the form of applications or packages to be directly installed on the devices themselves inducing local resource consumption. In that context, we have investigated a new approach for outsourcing mobile security functions as cloud-based services for smartphones and tablets [32]. The outsourced functions are dynamically activated, configured and orchestrated using software-defined networking and virtualization techniques. We consider the use of security compositions in order to dynamically fit the security requirements of mobile devices according to their current contexts. This approach is based on different traversal schemes (sequential, conditional, and concurrent). The solution has been prototyped based on the mininet software-defined networking emulator, jointly with mobile devices using the android operating system.

6.3.2. SDN-based security

Participants: Jérôme François [contact], Lautaro Dolberg [University of Luxembourg], Olivier Festor, Thomas Engel [University of Luxembourg].

By decoupling the data and control plane, Software-Defined Networking allows a fine grained network management. Protocols like OpenFlow allow multiple actions like traffic forwarding or blocking but also modifications or monitoring with the extensive use of counters. Hence, many approaches have emerged the last year to enable some security functions like firewalls, flow monitoring and traffic redirection to middleboxes. These different scenarios have been evaluated in a survey paper [17] in cooperation with the university of Luxembourg.

Furthermore, we also proposed to leverage SDN, especially OpenFlow, for forensics purpose [18]. Indeed, through a recursive analysis on network path and flow tables in OpenFlow, it is possible to reconstruct the paths traversing by an anomaly.

[0] http:///www.apisense.com/
[0] http://metroscope.eu/

6.3.3. Phishing Detection

Participants: Jérôme François [contact], Samuel Marchal [University of Luxembourg], Radu State [University of Luxembourg], Thomas Engel [University of Luxembourg].

This work is a joint work with the University of Luxembourg.

The language used for phishing is a particular language aiming at attracting victims. To achieve that the attackers uses specific words related to well known brand names and reassuring words. Our method to detect such abnormal domain names relies on word decomposition and semantic analysis. As an example, we can learn if having both *microsoft* and *protected* in domain is significative of a malicious domain. Actually, not all words can be represented during the learning and we use semantic similarities to also extend this knowledge (for example, we can *derive* safe from *protected*).

Our recent work [20] was focusing on extending this domain-based analysis to the full analysis of an url. We have also observed that most of false positives or negatives we obtained with previous methods are biased by natural language corpus while the *Internet vocabulary* is different.

Hence, we extracted from Google and Yahoo statistics about search queries. Our observation highlights that the relation between the different parts of the URL (the domain and the path) is a discriminative feature for malicious URL identification.

Finally, a more in-depth feature analysis is provided in [8], which also proposes leveraging streaming data analytics by instantiating our method on Storm.

6.3.4. Flows and logs analysis

Participants: Jérôme François [contact], Abdelkader Lahmadi.

Machine generated-log data is a fundamental part of information technology systems. They are usually generated at every component of distributed information systems including routers, security products, web proxies, DHCP servers, VPN servers, or any end-points like mobile devices or connected things, etc. They often contain high volumes of interesting information and are among the first data source to be analyzed for the detection of abnormal activities due to running attacks or malicious running applications. A better understanding of these attacks and malicious applications requires the elaboration of efficient and novel methods and techniques able to analyze these logs.

In [16], we carried an empirical analysis of the logs generated by the logging system available in Android environments. The logs are mainly related to the execution of the different components of applications and services running on an Android device. We have analyzed the logs using self organizing maps where our goal is to establish behavioral fingerprints of Android applications. The developed methodology allows us the better understand Android Apps regarding their granted permissions and performed actions.

During the year 2014, we have also maintained an IETF draft [50] to make a standardization effort towards the extension of IP Flow-based monitoring with geographic information. Associating Flow information with their measurement geographic locations will enable security applications to detect anomalous activities. In the case of mobile devices, the characterization of communication patterns using only time and volume is not enough to detect unusual location-related communication patterns.

6.3.5. Sensor networks monitoring

Participants: Rémi Badonnel, Isabelle Chrisment, Olivier Festor, Abdelkader Lahmadi [contact], Anthéa Mayzaud.

Low Power and Lossy Networks (LLNs) are made of interconnected wireless devices with limited resources in terms of energy, computing and communication. The communication channels are low-bandwidth, high loss rate and volatile wireless links subject to failure over time.

This year, our work on security-oriented monitoring [28] has focused on quantifying the effects of version number manipulation attacks within RPL networks [21]. Through simulations it was discovered that control overhead can increase by up to 18 times, thereby impacting energy consumption and channel availability. This in turn can reduce the delivery ratio of packets by up to 30% and nearly double the end-to-end delay in a network. A strong correlation between the position of the attacker and the effect on the network was also observed.

In that context, we have designed a mitigation strategy based on an adaptive threshold to cover a large variety of DODAG inconsistency attacks [25] in a lightweight manner. Currently RPL attempts to counteract such attacks by using a fixed threshold. During experimentations it becomes clear that the adaptive threshold is able to reduce the control message overhead, compared to fixed threshold, by up to 13% in short lived and 55% in long-lived networks. This leads to large reductions, i.e., between 10%-40%, in energy consumption.

In addition, we have investigated a distributed passive monitoring architecture for RPL-based advanced measurement infrastructure networks.

6.3.6. Intrusion Detection System in Wireless Sensor

Participants: Emmanuel Nataf [contact], Hubert Kenfack Ngankam.

This work is based on a previous work about the definition of an ontology to classify intrusion attacks in a wireless sensors network. A first implementation of this ontology focuses on the black hole and the sink hole intrusion where some malicious sensor node either do not forward data to a central point of collect or try to be elected as the best next hop toward the central point.

We look at discover malicious nodes by an analysis of the network topology obtained by data gathered from the network itself. At regular interval, we built a snapshot view of the network topology and compare it with the previous one in order to detect anomalies such as a whole sub network that disappear or an under-optimal network topology.

Simulation results are good and we will continue on this way.

6.3.7. SCADA systems security

Participants: Abdelkader Lahmadi [contact], Younes Abid.

SCADA systems are facing several attacks and threats which are growing in number and complexity. A key challenge in this context is the simulation and the assessment of the impact and the propagation of these attacks on SCADA system components over time. During the year 2014, we have developed a novel methodology [38] based on stochastic modeling to simulate the impact of attacks on SCADA systems. The system is modeled as a network of interacting markov chains and the impact of an attack is simulated using the influence model. In this model, the state of each node of the system is either influence by its own Markov chain or by the state of its neighboring nodes. We have modeled and analyzed a SCADA system with 200 control nodes and several servers. We have modeled different attacks (intrusion, DoS, malware) where attack nodes are introduced in the interacting SCADA network to influence control node behaviors. For each attack, we have simulated and assessed over time the availability of the overall system regarding the number of failed nodes.

6.3.8. Management of HTTPS traffic

Participants: Thibault Cholez [contact], Isabelle Chrisment, Shbair Wazen, Jérôme François.

Surveys show that websites are more and more being served over HTTPS. They highlight an increase of 48% of sites using TLS over the past year (2013),

We investigated the latest technique for HTTPS traffic filtering that is based on the Server Name Indication (SNI) field of TLS and which has been recently implemented in many firewall solutions. We show that SNI has two weaknesses, regarding (1) backward compatibility and (2) multiple services using a single certificate. We demonstrated thanks to a web browser plug-in called *Escape* that we designed and implemented, how these weaknesses can be practically used to bypass firewalls and monitoring systems relying on SNI. The results show positive evaluation (firewall's rules successfully bypassed) for all tested websites. This work will be published in the experience session of the IFIP/IEEE International Symposium on Integrated Network Management (IFIP/IEEE IM'15).

We also started a new work on the precise identification of websites accessed through HTTPS in the context of network forensic investigation. We use a new set of features in conjunction with machine learning techniques to achieve a high accuracy.

6.4. Routing

6.4.1. Routing in Wireless Sensor Networks

Participants: Emmanuel Nataf [contact], Patrick-Olivier Kamgueu.

We deployed a wireless sensors network in the laboratory during two time period of 3 months. The first was with the legacy routing (based on expected transmission time metric) and the second was with our routing process based on a composition of several metrics (i.e. energy, transmission time and delay) by the use of fuzzy logic. We have compared these experiments by packet loss ratio and energy consumption. In all case, our routing leads to a better network [48].

6.4.2. Operator calculus based routing in Wireless Sensor Networks

Participants: Evangelia Tsiontsiou, Bernardetta Addis, Ye-Qiong Song [contact].

For supporting different QoS requirements, routing in WSN must simultaneously consider several criteria (e.g., minimizing energy consumption, hop counts or delay, packet loss probability, etc.). When multiple routing metrics are considered, the problem becomes a multi-constrained optimal path problem (MCOP), which is known as NP-complete.

Recently, Operator calculus (OC) has been developed by Schott and Staples with whom we collaborate. We make use of OC methods on graphs to solve path selection in the presence of multiple constraints. Based on OC, we developed a distributed algorithm for path selection in a graph. We also designed a new routing protocol which makes use of this algorithm: the Operator Calculus based Routing Protocol (OCRP). In OCRP, a node selects the set of eligible next hops based on the given constraints and the distance to the destination. It then sends the packet to all eligible next hops. The protocol is implemented in Contiki OS and emulated for TelosB motes using Cooja. We compared its performance against tree and directional flooding routing and show the advantages of our technique. Our ongoing work consists in its comparison with RPL to show its effective contribution to handle simultaneously several IETF ROLL routing metrics.

This work is under development as part of Lorraine AME Satelor project.

6.4.3. Energy-aware IP networks management

Participants: Bernardetta Addis [contact], Giuliana Carello [DEIB, Politecnico di Milano, Italy], Antonio Capone [DEIB, Politecnico di Milano, Italy], Luca Gianoli [Polytecnique de Montreal, Canada], Sara Mattia [IASI, CNR, Roma, Italy], Brunide Sansò [Polytecnique de Montreal, Canada].

The focus of our research is to minimize the energy consumption of the network through a management strategy that selectively switches off devices according to the traffic level. We consider a set of traffic scenarios and jointly optimize their energy consumption assuming a per-flow routing. We propose a traffic engineering mathematical programming formulation based on integer linear programming that includes constraints on the changes of the device states and routing paths to limit the impact on quality of service and the signaling overhead. We also present heuristic results to compare the optimal operational planning with online energy management operation ([3])

Two very important issues that may be affected by green networking techniques are resilience to node and link failures, and robustness to traffic variations. We thus extended the optimization models. To guarantee network survivability we consider two different schemes, dedicated and shared protection, which assign a backup path to each traffic demand and some spare capacity on the links along the path. Robustness to traffic variations is provided by tuning the capacity margin on active links in order to accommodate load variations of different magnitude. Both exact and heuristic methods are proposed. Experimentations carried out on realistic networks operated with flow-based routing protocols (like MPLS) allow us to quantitatively analyze the trade-off between energy cost and level of protection and robustness. Results show that significant savings, up to 30%, may be achieved even when both survivability and robustness are fully guaranteed [4].

Computational cost of proposed models can be very high when dealing with large size instances (network size and/or number of demands). For this reason, we proposed and tested different problem formulations with the aim of solving larger size instances at optimality. Preliminary results on a simplified model ([29]) are very encouraging.

6.4.4. Energy-aware joint management of networks and Cloud infrastructures

Participants: Bernardetta Addis [contact], Danilo Ardagna [DEIB, Politecnico di Milano, Italy], Giuliana Carello [DEIB, Politecnico di Milano, Italy], Antonio Capone [DEIB, Politecnico di Milano, Italy].

Fueled by the massive adoption of Cloud services, overall service centers and networks account for 2–4% of global CO_2 emissions and it is expected they can reach up to 10% in 5–10 years.

The geographical distribution of the computing facilities offers many opportunities for optimizing energy consumption and costs by means of a clever distribution of the computational workload exploiting different availability of renewable energy sources, but also different time zones and hourly energy pricing. Energy and cost savings can be pursued by dynamically allocating computing resources to applications at a global level, while communication networks allow to assign flexibly load requests and to move data. We propose an optimization framework able to jointly manage the use of brown and green energy in an integrated system and to guarantee quality requirements. We propose an efficient and accurate problem formulation that can be solved for real-size instances in few minutes to optimality. Numerical results, on a set of randomly generated instances and a case study representative of a large Cloud provider, show that the availability of green energy have a big impact on optimal energy management policies and that the contribution of the network is far from being negligible ([2]).

6.4.5. Content centric wireless sensor networks

Participants: Abdelkader Lahmadi [contact], Younes Abid, Olivier Festor.

During this year, we have instantiated a novel named data aggregation method [9] dedicated to wireless sensor networks . The method relies on an adaptation of the CCNx protocol implementation that we have developed in a previous work. Our method extends the CCNx protocol with in-network processing functions to aggregate named data efficiently. We have implemented and tested our solution with the Contiki operating system which is an operating system for resources-constrained embedded systems and wireless sensor networks. Our simulation and measurement results using the Cooja simulator and physical nodes show that our solution has a small overhead in terms of exchanged messages and provides acceptable data retrieval delays.

6.5. Quality-of-Service

6.5.1. ICN cache management

Participants: Olivier Festor [contact], César Bernardini, Thomas Silverston.

Information Centric Networking (ICN) has become a promising new paradigm for the future Internet architecture. It is based on named data, where content address, content retrieval and the content identification is led by its name instead of its physical location. One of the ICN key concepts relies on in-network caching to store multiple copies of data in the network and serve future requests, which helps reducing the load on servers, congestion in the network and enhances end-users delivery performances. As a central component of ICN is in-network caching, the ely used as a micro-blogging service. At the same time, Online Social Networks (OSN) carry extremely valuable information about users and their relationships. We argue that this knowledge can help to drastically improve the efficiency of ICN.

We therefore propose SACS, a caching strategy designed for the CCN architecture that includes social information [11]. CCN is to date the most widely adopted ICN architecture by the research and industrial community. The underlying idea in such strategy is that a small number of users counts a huge amount of social relationships, dominates the activity and receives most attention from other users. We call such users Influential users, and we argue that they produce content that is more likely to be consumed by others, and in consequence their content must be favored and replicated in priority. Our novel caching strategy is therefore prioritizing content from Influential users of the social network. To validate our strategy, we first propose a model of social network over the CCN architecture [30]. Our model has been designed based on the measurement of Pinterest, a web-based OSN system. Extensive simulations of the strategy have been performed, as well as a real implementation on CCNx and deployment over the PlanetLab testbed. Our results with SACS are significant and increase drastically the caching performance of ICN architecture. content

Efficient management of caches is a key success factor in Concent-Centric Networks where multiple (up to every single node in the network) entities act as caches of the shared content in the network. We pursued our investigations towards a common evaluation framework for cache strategies in Content-centric networks and towards the definition of novel cache strategies, exploiting context information available at the service level of today's internet.

6.5.2. Self-adaptive MAC protocol for both QoS and energy efficiency

Participants: Kévin Roussel, Shuguo Zhuo, Ye-Qiong Song [contact].

WSN research focus has progressively been moved from the energy issue to the QoS issue. Typical example is the MAC protocol design, which cares about not only low duty-cycle at light traffic, but also high throughput with self-adaptation to dynamic traffic bursts.

The two MAC protocols that we have previously designed namely S-CoSenS and iQueue-MAC, have been successfully implemented on SMT32W108 SoC chips. Two contributions have been made this year. Firstly iQueue-MAC has been extended to work on both single channel mode and multi-channel mode, improving its throughput performance. Secondly, both S-CoSenS and iQueue-MAC have been implemented on RIOT OS. An additional contribution is related to the RIOT OS development itself since we have improved the robustness of the existing ports of RIOT OS on MSP430-based motes, making it a suitable software platform for tiny motes and devices. More generally, through this part of work, we have shown that RIOT OS is also suitable for implementing high-performance MAC protocols, thanks to its real-time features (especially hardware timers management). Part of this work has been supported by ANR-NFSC Quasimodo and PIA LAR projects.

6.5.3. End-to-end delay modelling and evaluation in wireless sensor networks

Participants: François Despaux, Abdelkader Lahmadi, Ye-Qiong Song [contact].

Probabilistic end-to-end performance guarantee may be required when dealing with real-time applications. As part of ANR QUASIMODO project, we are dealing with Markov modeling of multi-hop networks running duty-cycled MAC protocols. One of the problems of the existing Markovian models resides in their strong assumptions that may not be directly used to assess the end-to-end delay in practice. In particular, realistic radio channel, capture effect and OS-related implementation factors are not taken into account. We proposed to explore a new approach combining code instrumentation and Markov chain analysis. In [15] we have presented a new approach for extracting empirical Markov chain models from network protocol traces by means of Process Mining techniques. An empirical Markov chain model was obtained for the IEEE 802.15.4 beacon-enabled mode protocol allowing us to estimate the e2e delay for a multi-hop scenario. This approach has also been successfully applied to the case of ContikiMAC [14].

6.5.4. Dynamic resource allocation in network virtualization

Participants: Mohamed Said Seddiki, Mounir Frikha [SupCom, Tunis, Tunisie], Ye-Qiong Song [contact].

The objective of this research topic is to develop different resource allocation mechanisms in Network Virtualization, for creating multiple virtual networks (VNs) from a single physical network. It is accomplished by logical segmentation of the network nodes and their physical links.

This year we have focused on implementing and evaluating the used of SDN for managing the QoS in broadband access networks. Unfortunately, application-based QoS on a home network gateway faces significant constraints, as commodity home routers are not typically powerful enough to perform application classification, and many home users are not savvy enough to configure QoS parameters. In [24] we designed FlowQoS, an SDN-based approach where users can specify upstream and downstream bandwidth allocations for different applications at a high level, offloading application identification to an SDN controller that dynamically installs traffic shaping rules for application flows We designed a custom DNS-based classifier to identify different applications that run over common web ports; a second classifier performs lightweight packet inspection to classify non-HTTP traffic flows. We implemented FlowQoS on OpenWrt and demonstrated that it can improve the performance of both adaptive video streaming and VoIP in the presence of active competing traffic.

This work has been carried out as part of a co-supervised PhD thesis between University of Lorraine and SupCom Tunis.

6.5.5. Task and message scheduling in distributed real-time systems

Participants: Florian Greff, Laurent Ciarletta, Ye-Qiong Song [contact].

QoS must be guaranteed when dealing with real-time distributed systems interconnected by a network. Not only task schedulability in processors, but also message schedulability in networks should be analysed for validating the system design. In [37], [36], [34], and [35], we provided an overview of both message scheduling techniques in networks and joint task and message scheduling approaches in closed-loop distributed control systems (networked control systems). Fault-tolerance is another critical issue that one must take into account. In collaboration with an industrial partner, we started a study on the real-time dependability of UAV multi-criticity system interconnected by an embedded mesh network. The future work aims at developing a robust mesh network routing protocol and studying the schedulability under constraints of multi-criticality and graceful degradation during mode change.

6.6. Multi-modeling and co-simulation tools for the evaluation and development of Smart* and other Pervasive Computing systems

Participants: Laurent Ciarletta [contact], Olivier Festor, Ye-Qiong Song, Yannick Presse, Emmanuel Nataf, Benjamin Segault.

Vincent Chevrier (Maia team, LORIA) is a collaborator and the correspondant for the MS4SG project, Benjamin Camus, Victorien Elvinger and Christine Bourjot (Maia team, LORIA) are collaborators for the AA4MM. Julien Vaubourg's PhD is under the co-direction of V. Chevrier and L. Ciarletta.

In Pervasive or Ubiquitous Computing, a growing number of communicating/computing devices are collaborating to provide users with enhanced and ubiquitous services in a seamless way.

These systems, embedded in the fabric of our daily lives, are complex: numerous interconnected and heterogeneous entities are exhibiting a global behavior impossible to forecast by merely observing individual properties. Firstly, users physical interactions and behaviors have to be considered. They are influenced and influence the environment. Secondly, the potential multiplicity and heterogeneity of devices, services, communication protocols, and the constant mobility and reorganization also need to be addressed. Our research on this field is going towards both closing the loop between humans and systems, physical and computing systems, and taming the complexity, using multi-modeling (to combine the best of each domain specific model) and co-simulation (to design, develop and evaluate) as part of a global conceptual and practical toolbox.

We proposed the AA4MM meta-model [51] that solves the core challenges of multimodeling and simulation coupling in an homogeneous perspective. In AA4MM, we chose a multi-agent point of view: a multi-model is a society of models; each model corresponds to an agent and coupling relationships correspond to interaction between agents. In the MS4SG projet which involves MAIA, Madynes and EDF R&D on smart-grid simulation, we developed a proof of concepts for a smart-appartment case[12].

In 2014 we worked on the following research topics:

- Assessment and evaluation of complex systems.

  This work, centered on the problem of controlling complex systems proposed a control architecture within Tomas Navarrete's work [22], [23]. This "equation-free" approach uses a multi-agent model to evaluate the global impact of local control actions before applying the most pertinent set of actions. Based on a partial perception of the system state, we determine which actions to execute in order to avoid or favor certain global states of the system.

  Associated to our architecture, an experimental platform has been developed to confront the basic ideas or the architecture within the context of simulated "free-riding" phenomenon in peer to peer file exchange networks. We have demonstrated that our approach allows us to drive the system to a state where most peers share files, despite given initial conditions that are supposed to drive the system to a state where no peer shares.

- Cyber Physical Systems [13]

  We have led the design and implementation of the Aetournos platform at Loria. The collective movements of a flock of flying communicating robots / UAVs, evolving in potentially perturbed environment constitute a good example of a Cyber Physical System. Applying co-simulation technique we plan to develop a hybrid "network-aware flocking behavior" / "behavior aware routing protocol".

  We have provided a working set of tools: multi-simulation behavior / network / physics and generic software development using ROS (Robot Operating System). The UAVs carry a set of sensor for location awareness, their own computing capabilities and several wireless networks.

  The effort put in the UAVs gathers academic and research ressources from the Aetournos platform, the R2D2 ADT and the 6PO project, while applied, industrial and more R&D projects have been pursued this year (Outback Joe Search and Rescue Challenge, Alerion, Hydradrone) .

- MS4SG has given us the opportunity to link multi-simulations tools such as HLA (High Level Architecture) and FMI (Functional Mockup Interface) thanks to our AA4MM framework. We have so far successfully applied our solution to the simulation of smart apartment complex and to combine the electrical and networking part of a Smart Grid[12].

In 2015, we will continue working on the hybrid protocols and on the UAV platform, and apply our co-simulation work to Smart Grids and other Smart*.

MAESTRO Project-Team

6. New Results

6.1. Highlights of the Year

E. Altman has received the "Isaacs' Award" granted by the International Society on Dynamic Games in recognition for his research on dynamic game theory.

M. El Chamie got the Best Session Presentation Award at the IEEE American Control Conference ACC 2014 for the paper "Newton's method for constrained norm minimization and its application to weighted graph problems," co-authored with G. Neglia.

THANES is a new French-Brazilian joint-team between MAESTRO and researchers from Univ. Federal do Rio de Janeiro (Brazil) and Carneggie Mellon Univ. (USA). The team investigates network science problems with a particular focus on Online Social Networks.
BEST PAPERS AWARDS :
[43] **6th IEEE INFOCOM International Workshop on Network Science for Communication Networks (NetSciCom)**. K. AVRACHENKOV, P. BASU, G. NEGLIA, B. RIBEIRO, D. TOWSLEY.
[70] **4th IEEE Online Conference on Green Communications (GreenComm)**. C. ROTTONDI, G. NEGLIA, G. VERTICALE.

6.2. Network Science

Participants: Eitan Altman, Konstantin Avrachenkov, Mahmoud El Chamie, Julien Gaillard, Arun Kadavankandy, Jithin Kazhuthuveettil Sreedharan, Hlib Mykhailenko, Philippe Nain, Giovanni Neglia, Yonathan Portilla, Alexandre Reiffers, Vikas Singh, Marina Sokol.

6.2.1. Epidemic models of propagation of content

Epidemic models have received significant attention in the past few decades to study the propagation of viruses, worms and ideas in computer and social networks. In the case of viruses, the goal is to understand how the topology of the network and the properties of its nodes impact the spread of the epidemics. In [38], E. Altman, A. Avritzer and L. Pfleger de Aguiar (Siemens Corporation, Princeton, USA), R. El-Azouzi (Univ. of Avignon), and D. S. Menasche (Federal Univ. of Rio de Janeiro, Brazil) propose rejuvenation as a way to cope with epidemics. Reformatting a computer may solve the problem of virus contamination (but it might be a costly operation) while less dramatic actions may render the computer operational again (even in the presence of the virus). In this work they evaluate the performance gain of such measures as well as sampling for early detection of viruses while these incubate. During incubation, contaminated terminals are infectious and yet, if not detected to be so, they cannot be isolated and treated.

In [60], Y. Hayel (Univ. of Avignon), S. Trajanovski and P. Van Mieghem (Delft Univ. of Technology, The Netherlands), E. Altman, and H. Wang (Delft Institute of Applied Mathematics, The Netherlands), compare solutions involving vaccination to those that involve healing from a selfish point of view of an individual networked user. A game theoretical model is presented and the obtained equilibrium is computed for various types of topologies including the fully connected one, the bipartite graph and a community structure. A novel use of potential games is presented to compute the equilibria.

In [61], L. Maggi and F. De Pellegrini (CREATE-NET, Italy), A. Reiffers, J. J. Herings (Maastricht Univ., The Netherlands) and E. Altman, study a viral diffusion of a content in a multi-community environment. Exploiting time scale separation, the authors are able to reduce the dimensionality of the problem and to compute its limiting behavior in closed form. They further study regulation and cooperative approaches for sharing the cost for fighting the spread of the infection among the communities.

Social networks can have asymmetric relationships. In the online social network Twitter, a follower receives tweets from a followed person but the followed person is not obliged to subscribe to the channel of the follower. Thus, it is natural to consider the dissemination of information in directed networks. In [44], K. Avrachenkov in collaboration with B. Prabhu (LAAS-CNRS), K. De Turck and D. Fiems (Ghent Univ., Belgium) use the mean-field approach to derive differential equations that describe the dissemination of information in a social network with asymmetric relationships. In particular, their model reflects the impact of the degree distribution on the information propagation process. They further show that for an important subclass of their model, the differential equations can be solved analytically.

6.2.2. Bio-Inspired Models for Characterizing YouTube Viewcount

Bio-inspired models have long been advocated for the dissemination of content in the Internet. How good are such models and how representative are they? In [69], C. Richier, R. El-Azouzi, T. Jimenez, G. Linares (all with Univ. of Avignon), E. Altman and Y. Portilla propose six different epidemic models. These are classified according to various criteria: (i) the size of the target population, which may be constant, or linearly increasing or infinite, (ii) the virality of the content: it is said to be viral if nodes that receive the content participate in retransmitting it (by sharing or embedding). They then collected data on the viewcounts of videos in youtube and examined how well they fit their models. They showed that their six models cover 90% of the videos with an average mean square error of less than 5%. They further studied the capability of using these models to predict the evolution of the viewcount.

6.2.3. Network centrality measures

Finding quickly top-k lists of nodes with the largest degrees in large complex networks is a basic problem of recommendation systems. If the adjacency list of the network is known (not often the case in complex networks), a deterministic algorithm to solve this problem requires an average complexity of $O(n)$, where n is the number of nodes in the network. Even this modest complexity can be excessive for large complex networks. In [18], K. Avrachenkov and M. Sokol in collaboration with N. Litvak (Twente Univ., The Netherlands) and D. Towsley (Univ. of Massachusetts, Amherst, USA) propose to use a random-walk-based method. They show theoretically and by numerical experiments that for large networks, the random-walk method finds good-quality top lists of nodes with high probability and with computational savings of orders of magnitude. They also propose stopping criteria for the random-walk method that requires very little knowledge about the structure of the network.

In [46], K. Avrachenkov in collaboration with N. Litvak (Twente Univ., the Netherlands) and L. Ostroumova and E. Suyargulova (both from Yandex, Russia) address the problem of quick detection of high-degree entities in large online social networks. The practical importance of this problem is attested by a large number of companies that continuously collect and update statistics about popular entities, usually using the degree of an entity as an approximation of its popularity. They suggest a simple, efficient, and easy to implement two-stage randomized algorithm that provides highly accurate solutions for this problem. For instance, their algorithm needs only one thousand API requests in order to find the top-100 most followed users in Twitter, a network with approximately a billion of registered users, with more than 90% precision. Their algorithm significantly outperforms existing methods and serves many different purposes, such as finding the most popular users or the most popular interest groups in social networks. They show that the complexity of the algorithm is sublinear in the network size, and that high efficiency is achieved in networks with high variability among the entities, expressed through heavy-tailed distributions.

Personalized PageRank is an algorithm to classify the importance of web pages on a user-dependent basis. In [48], K. Avrachenkov and M. Sokol in collaboration with R. van der Hofstad (EURANDOM, The Netherlands) introduce two generalizations of Personalized PageRank with node-dependent restart. The first generalization is based on the proportion of visits to nodes before the restart, whereas the second generalization is based on the proportion of time a node is visited just before the restart. In the original case of constant restart probability, the two measures coincide. They discuss interesting particular cases of restart probabilities and restart distributions. They show that both generalizations of Personalized PageRank have an elegant expression connecting the so-called direct and reverse Personalized PageRanks that yield a symmetry property of these Personalized PageRanks.

Along with K. Avrachenkov and N. M. Markovich (Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia), J. K. Sreedharan investigated distribution and dependence of extremes in network sampling processes [47]. This is one of the first studies associating extremal value theory to sampling of large networks. The work showed that for any general stationary samples from the graph (function of node samples) meeting two mixing conditions, the knowledge of bivariate distribution or bivariate copula is sufficient to derive many of its extremal properties. The work proved the usage of a single parameter to find many relevant extremes in networks like order statistics, first hitting time, mean cluster size etc. In particular, correlation in degrees of adjacent nodes are modelled and different random walks, such as PageRank, are studied in detail. This work has been done in the context of Inria Alcatel-Lucent Bell Labs joint laboratory's ADR "Network Science" (see §7.1.2).

6.2.4. Influence maximization in complex networks

Efficient marketing or awareness-raising campaigns seek to recruit a small number, w, of influential individuals—where w is the campaign budget—that are able to cover the largest possible target audience through their social connections. In [43] K. Avrachenkov and G. Neglia in collaboration with P. Basu (BBN Technologies, US), B. Ribeiro (CMU, US) and D. Towsley (Univ. of Massachusetts, Amherst, USA) assume that the topology is gradually discovered thanks to recruited individuals disclosing their social connections. They analyze the performance of a variety of online myopic algorithms (i.e. that do not have a priori information on the topology) currently used to sample and search large networks. They also propose a new greedy online algorithm, Maximum Expected Uncovered Degree (MEUD). Their proposed algorithm greedily maximizes the expected size of the cover, but it requires the degree distribution to be known. For a class of random power law networks they show that MEUD simplifies into a straightforward procedure, denoted as MOD because it requires only the knowledge of the Maximum Observed Degree. This work has been done in the context of THANES Joint team (see §8.3.1.1) and Inria Alcatel-Lucent Bell Labs joint laboratory's ADR "Network Science" (see §7.1.2).

In [66] G. Neglia, in collaboration with X. Ye (Politecnico di Torino, Italy), M. Gabielkov and A. Legout (from the DIANA team) consider how to maximize users influence in Online Social Networks (OSNs) by exploiting social relationships only. Their first contribution is to extend to OSNs the model of Kempe, Kleinberg and Tardös on the propagation of information in a social network and to show that a greedy algorithm is a good approximation of the optimal algorithm that is NP- hard. However, the greedy algorithm requires global knowledge, which is hardly practical. Their second contribution is to show on simulations on the full Twitter social graph that simple and practical strategies perform close to the greedy algorithm.

6.2.5. Clustering

Clustering of a graph is the task of grouping its nodes in such a way that the nodes within the same cluster are well connected, but they are less connected to nodes in different clusters. In [45] K. Avrachenkov, M. El Chamie and G. Neglia propose a clustering metric based on the random walks' properties to evaluate the quality of a graph clustering. They also propose a randomized algorithm that identifies a locally optimal clustering of the graph according to the metric defined. The algorithm is intrinsically distributed and asynchronous. If the graph represents an actual network where nodes have computing capabilities, each node can determine its own cluster relying only on local communications. They show that the size of clusters can be adapted to the available processing capabilities to reduce the algorithm's complexity.

6.2.6. Average consensus protocols

In [54], [82], M. El Chamie in collaboration with J. Liu and T. Başar (Univ. of Illinois at Urbana Champaign, USA) studies the performance of a subclass of distributed averaging algorithms where the information exchanged between neighboring nodes (agents) is subject to deterministic uniform quantization. They give the convergence properties of linear averaging due to such quantization (which is a practical concern for many applications) that cause nonlinearity in the system. This is the first attempt to solve the exact model.

In [53], M. El Chamie in collaboration with T. Başar (Univ. of Illinois at Urbana Champaign, USA) considers optimal design strategies in consensus protocols for networks vulnerable to adversarial attacks. They provide a game theoretical model for the problem of a network with an adversary corrupting the control signal with noise. They derive the optimal strategies for both players (the adversary and the network designer) of the resulting game using a saddle point equilibrium solution in mixed strategies.

6.3. Wireless Networks

Participants: Eitan Altman, Abdulhalim Dandoush, Majed Haddad, Jithin Kazhuthuveettil Sreedharan.

6.3.1. Localization in ad-hoc wireless sensors networks

Range-based localization algorithms in wireless sensor networks are more accurate but also more computationally complex than the range-free algorithms. In collaboration with M. S. Elgamel (Univ. of Louisiana, USA), A. Dandoush has revised the Trigonometric based Ad-hoc Localization System (TALS) proposed in the literature. In [83], they propose a new technique to optimize the system: by eliminating the need of solving a linear system of equations via least square methods or its variants or the need for any square root operations, the computational overhead is reduced. Also, a novel modified Manhattan distance is proposed and used in the elimination process ensuring thereby a very good accuracy with less complexity than the basic TALS. Through a mathematical analysis and intensive simulations, the optimized TALS is shown to present superior performance and accuracy results compared to other localization techniques.

6.3.2. Channel management

The enhanced Inter Cell Interference Coordination (eICIC) feature has been introduced to solve the interference problem in small cells. It involves two parameters which need to be optimized, namely the Cell Range Extension (CRE) of the small cells and the ABS ratio (ABSr) which defines a mute ratio for the macro cell to reduce the interference it produces. In [72], A. Tall, Z. Altman (Orange Labs, Issy les Moulineaux) and E. Altman propose self-optimizing algorithms for the eICIC. The CRE is adjusted by means of a load balancing algorithm. The ABSr parameter is optimized by maximizing a proportional fair utility of user throughputs. The convergence of the algorithms is proven using Stochastic Approximation theorems. Numerical simulations illustrate the important performance gain brought about by the different algorithms.

Cognitive Radios are proposed as a solution to scarcity of wireless spectrum and one of the main challenges here is to gain knowledge about the spectrum usage by the licensed users, termed as spectrum sensing. In [29], Vinod Sharma (Indian Institute of Science, Bangalore, India) and J. K. Sreedharan study novel algorithms for spectrum sensing which minimize the expected time for spectrum sensing with stringent constraints on the probability of wrong detection. Algorithms are distributed in nature and the work proves that the algorithms are asymptotically optimal distributed sequential hypothesis tests. Along with theoretical guarantees, many practical scenarios in Cognitive Radios are also investigated.

6.3.3. Self-Organizing Network (SON)

The fast development of SON technology in mobile networks renders critical the problem of coordinating SON functionalities operating simultaneously. SON functionalities can be viewed as control loops that may need to be coordinated to guarantee conflict free operation, to enforce stability of the network and to achieve performance gain. In [30], A. Tall and Z. Altman (Orange Labs, Issy les Moulineaux), R. Combes (SUPELEC), and E. Altman propose a distributed solution for coordinating SON functionalities. It uses Rosen's concave games framework in conjunction with convex optimization. The SON functionalities are modeled as linear Ordinary Differential Equation (ODE)s. The stability of the system is first evaluated using a basic control theory approach together with strict diagonal concavity notion that originates from game theory. The coordination solution consists in finding a linear map (called coordination matrix) that stabilizes the system of SON functionalities. It is proven that the solution remains valid in a noisy environment using Stochastic Approximation.

6.4. Network Engineering Games

Participants: Eitan Altman, Ilaria Brunetti, Majed Haddad, Alexandre Reiffers.

6.4.1. The association problem

In [57], M. Haddad, S. Habib (Orange Labs, Issy les Moulineaux), and P. Wiecek (Wroclaw Univ. of Technology, Poland) and E. Altman develop a hierarchical Bayesian game framework for automated dynamic offset selection. Users compete to maximize their throughput by picking the best locally serving radio access network (RAN) with respect to their own measurement, their demand and a partial statistical channel state information of other users. In particular, they investigate the properties of a Stackelberg game, in which the base station is a player on its own. They derive analytically the utilities related to the channel quality perceived by users to obtain the equilibria. They study the Price of Anarchy of such system, which is defined as the ratio of the social welfare attained when a network planner chooses policies to maximize social welfare versus the social welfare attained at a Nash/Stackeleberg equilibrium when users choose their policies strategically.

6.4.2. Cognitive radio

In [26], M. Haddad, P. Wiecek (Wroclaw Univ. of Technology, Poland), O. Habachi and Y. Hayel (both with Univ. of Avignon) propose a game theoretical approach that allows cognitive radio pairs, namely the primary user (PU) and the secondary user (SU), to update their transmission powers and frequencies simultaneously. Specifically, a Stackelberg game model in which individual users attempt to hierarchically access to the wireless spectrum while maximizing their energy efficiency was addressed. A thorough analysis of the existence, uniqueness and characterization of the Stackelberg equilibrium was conducted. In particular, it was shown that a spectrum coordination naturally occurs when both actors in the system decide sequentially about their powers and their transmitting carriers. As a result, spectrum sensing in such a situation turns out to be a simple detection of the presence/absence of a transmission on each sub-band. An algorithmic analysis on how the PU and the SU can reach such a spectrum coordination using an appropriate learning process is provided.

In [59], the same authors present a hierarchical game to model distributed joint power and channel allocation for multi-carrier energy efficient cognitive radio systems. A thorough analysis of the existence, uniqueness and characterization of the Stackelberg equilibrium is conducted. It was proved that, at the Stackelberg equilibrium, each of the two users transmits on only one carrier depending on the fading channel gains. This results contrast with capacity-based approaches in which a certain number of carriers is exploited depending on the channel gains. Interestingly, it was shown that, for the vast majority of cases, introducing a certain degree of hierarchy in a multi-carrier system induces a natural coordination pattern where users have incentive to choose their transmitting carriers in such a way that they always transmit on orthogonal channels. Analytical results were provided for assessing and improving the performances in terms of energy efficiency between the non-cooperative game with synchronous decision makers and the proposed Stackelberg game.

6.4.3. Routing Games

In [39], E. Altman, J. Kuri (Indian Institute of Science, Bangalore, India) and R. El-Azouzi (Univ. of Avignon) study a routing game that models competition over a simple network with losses. Packets may be lost in the network due to either congestion losses or to channel random losses. They compute the equilibrium and establish its properties. They identify a Braess type paradox in which by adding a link the loss probabilities of all players increase.

G. Accongiagioco (Institute for Advanced Studies, Lucca, Italy), E. Altman, E. Gregori (Italian National Research Council, Italy) and L. Lenzini (Univ. of Pisa, Italy) analyze in [36] the decisions taken by an Autonomous System (AS) when joining the Internet. They first define a realistic model for the interconnection costs incurred and then they use this cost model to perform a game theoretic analysis of the decisions related to the creation of new links in the Internet. The proposed model does not fall into the standard category of routing games, hence they devise new tools to solve it by exploiting peculiar properties of the game. They prove analytically the existence of multiple equilibria for specific cases, and provide an algorithm to compute the stable ones. The analysis of the model's outcome highlights the existence of a Price of Anarchy and a Price of Stability.

6.4.4. Network neutrality and collusion

Representatives of several Internet access providers have expressed their wish to see a substantial change in the pricing policies of the Internet. In particular, they would like to see content providers pay for use of the network, given the large amount of resources they use. This would be in clear violation of the "network neutrality" principle that had characterized the development of the wireline Internet. In [14], E. Altman, M. K. Hanawal (former PhD student in MAESTRO) and R. Sundaresan (Indian Institute of Science, Bangalore, India) proposed and studied possible ways of implementing such payments and of regulating their amount. The results were reported already in a previous report, but were substantially revised during the period of this project.

6.4.5. Competition over popularity in social networks

We have pursued our analysis of competition over popularity and visibility in social networks. In [68], A. Reiffers and E. Altman, together with Y. Hayel (Univ. of Avignon) study a game model that arises when the rate of transmission of packets of each source can be accelerated in order to optimize a weighted sum of its acceleration cost and the expected number of its contents on the timelines of those who follow that content. While this paper considers equilibrium within static policies (in which the acceleration rate does not change in time), the same authors study in [51] the structure of dynamic equilibrium policies which are allowed to change as a function of the time (or of the state). A problem with a similar tradeoff is studied by E. Altman in a mobile context in [13] where the question of accelerating the transmission rate of content arises in a context of competition over content where it is assumed that if a content reaches a given destination then that destination will not be interested any more in receiving competing content.

In [67], A. Reiffers, E. Altman and Y. Hayel (Univ. of Avignon) extend the work in [68], and model the situation in which several social networks are available and a source may control not only the rate of transmission (acceleration) but may also decide how to split its content to the various social networks.

A competition over the timing of the transmission of a content was studied by E. Altman and N. Shimkin (Israel Institute of Technology, Israel) in [41]. Uniqueness of a symmetric equilibrium was established under the assumption of Poisson arrival of requests.

6.5. Green Networking and Smart Grids

Participants: Sara Alouf, Eitan Altman, Alberto Benegiamo, Ioannis Dimitriou, Majed Haddad, Alain Jean-Marie, Giovanni Neglia.

6.5.1. Energy efficiency in wireless networks

In [25], M. Haddad, P. Wiecek (Wroclaw Univ. of Technology, Poland), O. Habachi and Y. Hayel (both with Univ. of Avignon) investigated the achievable performances of multi-carrier energy efficient power control game. Both the simultaneous-move and the hierarchical games were addressed. For the first time, the analytical closed-form expressions of the spectrum coordination and the spectral efficiency of such models was derived. Results indicate that the spectrum coordination capability induced by the power control game model enables the wireless network to enjoy the energy efficiency improvement while still achieving a high spectral efficiency.

In [58], the same authors studied energy efficiency of heterogeneous networks for both sparse and dense (two-tier and multi-tier) small cell deployments. The problem is formulated as a hierarchical (Stackelberg) game in which the macro cell is the leader whereas the small cell is the follower. Both players want to strategically decide on their power allocation policies in order to maximize the energy efficiency of their registered users. A backward induction method has been used to obtain a closed-form expression of the Stackelberg equilibrium. It was shown that the energy efficiency is maximized when only one sub-band is exploited for the players of the game depending on their fading channel gains.

In [34], R. A. Vaca Ramirez and J. S. Thompson (Univ. of Edinburgh, UK), E. Altman and V. M. Ramos Ramos (Univ. Autonoma Metropolitana, Mexico) aim to reduce the power expenditure in the reverse link during low network load periods, by allocating extra resource blocks (RBs) to the mobile users. This is in contrast with other approaches in which resources are reduced in hours of low energy consumption. The user's rate demands are split among its allocated RBs in order to transmit in each of them by using a low level modulation order. In this low SINR regime the transmission is much more energy efficient since the log appearing in Shannon formula is in close to linear. We model the bandwidth expansion (BE) process by a game theory framework derived from the concept of stable marriage with incomplete lists (SMI).

P. Wiecek (Wroclaw Univ. of Technology, Poland) and E. Altman consider in [42] dynamic Multiple Access games between a random number of players competing over collision channels. Each of several mobiles involved in an interaction determines whether to transmit at a high or at a low power. High power decreases the lifetime of the battery but results in smaller collision probability. They formulated this game as an anonymous sequential game with undiscounted reward and computed the equilibrium [42]. The internal state of a player corresponds to the amount of energy left in the battery and the actions correspond to the transmission power.

I. Dimitriou investigated in [52] the power management of mobile devices, using a variant of an M/G/1 queue with probabilistic inhomogeneous multiple vacations and generalized service process. Under the vacation scheme, at the end of a vacation the server goes on another vacation, with a different probability distribution, if during the previous vacation there have been no arrivals. The modified vacation policy depends on the initial vacation interval and the server selects randomly over M such vacation policies. The theoretical system can be applied for modeling the power saving mode of mobile devices in modern wireless systems. Moreover, the form of the service process properly describes the incremental redundancy retransmission scheme that provides different types of retransmissions in such systems. Steady state analysis is investigated, energy and performance metrics are obtained and used to provide numerical results that are also validated against simulations.

6.5.2. Energy efficiency in delay tolerant networks

Energy efficiency in mobile networks is further studied in [28] where L. Sassatelli (Univ. of Nice Sophia Antipolis), A. Ali, M. Panda and T. Chahed (all with Telecom SudParis) and E. Altman tackle the issue of reliable transport in Delay-Tolerant mobile ad hoc Networks, that are operated by some opportunistic routing algorithm. We propose a reliable transport mechanism that relies on Acknowledgements (ACK) and coding at the source. The various versions of the problem depending on buffer management policies are formulated, and a fluid model based on a mean-field approximation is derived for the designed reliable transport mechanism. This model allows to express both the mean file completion time and the energy consumption up to the delivery of the last ACK at the source.

6.5.3. Modeling of a smart green base station

S. Alouf, I. Dimitriou A. Jean-Marie have considered the modeling of wireless communication base stations with autonomous energy supply (solar, wind). They proposed and analyzed a queueing model to assess performance of a base station fully powered by renewable energy sources. The system operates in a finite state Markovian random environment that properly describes the intermittent nature of renewable energy sources and the data traffic. The base station is considered to be "smart" in the sense that it is able to dynamically adjust its coverage area, controlling thereby the traffic rate and its energy consumption. They show how the matrix-analytic formalism enables to construct and study the performance of a smart green base station operating in random environment. More precisely, the behavior of such a system is described by a five-dimensional Markov process, which is a homogeneous finite Quasi Birth-Death (QBD) process. Several existing algorithms can be used in order to obtain the stationary probability vector, which is the basis for the calculation of interesting performance metrics. This work is on-going and has not been submitted for publication yet.

6.5.4. Direct Load Control

Balancing energy demand and production is becoming a more and more challenging task for energy utilities also because of the larger penetration of renewable energies which are more difficult to predict and control.

While the traditional solution is to dynamically adapt energy production to follow the time-varying demand, a new trend is to drive the demand itself. Most of the ongoing actions in this direction involve greedy energy consumers, like industrial plant, supermarkets or large buildings. Pervasive communication technologies may allow in the near future to push further the granularity of such approach, by having the energy utility interacting with residential appliances. In [65] and in its extension [64], G. Neglia, in collaboration with G. Di Bella, L. Giarré and I. Tinnirello (Univ. of Palermo, Italy) study large scale direct control of inelastic home appliances whose energy demand cannot be shaped, but simply deferred. Their solution does not suppose any particular intelligence at the appliances. The actuators are rather smart plugs (simple devices with local communication capabilities that can be inserted between appliances plugs and power sockets) and are able to interrupt/reactivate power flow through the plug. A simple control message can be broadcast to a large set of smart plugs for probabilistically enabling or deferring the activation requests of a specific load type in order to satisfy a probabilistic bound on the aggregated power consumption. The control law and the most important performance metrics can be easily derived analytically.

6.5.5. Charge of Electric Vehicles

The massive introduction of Electric Vehicles (EVs) is expected to significantly increase the power load experienced by the electrical grid, but also to foster the exploitation of renewable energy sources: if the charge process of a fleet of EVs is scheduled by an intelligent entity such as a load aggregator, the EVs' batteries can contribute in flattening energy production peaks due to the intermittent production patterns of renewables by being recharged when energy production surpluses occur. To this aim, time varying energy prices are used, which can be diminished in case of excessive energy production to incentivize energy consumption (or increased in case of shortage to discourage energy utilization). In [70] G. Neglia, in cooperation with C. Rottondi and G. Verticale (Politecnico di Milano, Italy), evaluate the complexity of the optimal scheduling problem for a fleet of EVs aimed at minimizing the overall cost of the battery recharge in presence of time-variable energy tariffs. The scenario under consideration is a fleet owner having full knowledge of customers' traveling needs at the beginning of the scheduling horizon. They prove that the problem has polynomial complexity, provide complexity lower and upper bounds, and compare its performance to a benchmark approach which does not rely on prior knowledge of customers' requests, in order to evaluate whether the additional complexity required by the optimal scheduling strategy w.r.t. the benchmark is worthy the achieved economic advantages. Numerical results show considerable cost savings obtained by the optimal scheduling strategy.

6.6. Content-Oriented Systems

Participants: Sara Alouf, Eitan Altman, Konstantin Avrachenkov, Nicaise Choungmo Fofack, Abdulhalim Dandoush, Majed Haddad, Alain Jean-Marie, Philippe Nain, Giovanni Neglia, Marina Sokol.

6.6.1. Modeling modern DNS caches

N. Choungmo Fofack and S. Alouf have pursued their study of the modern behavior of DNS (Domain Name System) caches. The entire set of traces collected in 2013 by Inria's IT services in Sophia Antipolis at one of the Inria's DNS caches have been processed and analyzed with the help of N. Nedkov (4-month intern in MAESTRO). This allowed to strengthen the validation of the theoretical models developed in 2013 (see [86]). On the other hand, parts of [86] have been revisited and derived under more general assumptions. As a direct consequence, the exact analysis derived on linear cache networks is extended to a large class of hierarchical cache networks called *linear-star* networks which include linear and two-level tree/star networks. In addition, closed-form expressions for the cache consistency measures (refresh rate and correctness probability) are provided under the assumption that contents requests and updates occur according to two independent renewal processes.

6.6.2. Analysis of general and heterogeneous cache networks

There has been considerable research on the performance analysis of *on-demand* caching replacement policies like Least-Recently-Used (LRU), First-In-First-Out (FIFO) or Random (RND). Much progress has been made on the analysis of a single cache running these algorithms. However it has been almost impossible to extend the results to networks of caches. In [22], N. Choungmo Fofack, P. Nain and G. Neglia, in collaboration with D. Towsley (Univ. of Massachusetts, Amherst, USA), introduce a Time-To-Live (TTL) based caching model, that assigns a timer to each content stored in the cache and redraws it every time the content is requested (at each hit/miss). They derive the performance metrics (hit/miss ratio and rate, occupancy) of a TTL-based cache in isolation fed by stationary and ergodic request processes with general TTL distributions. Moreover they propose an iterative procedure to analyze TTL-based cache networks under the assumptions that requests are described by *renewal processes* (that generalize Poisson processes or the standard IRM assumption). They validate the theoretical findings through event-driven and Monte-Carlo simulations based on the Fourier Amplitude Sensitivity Test to explore the space of the input parameters. The analytic model predicts remarkably well all metrics of interest with relative errors smaller than 1%.

Jointly with M. Badov, M. Dehghan, D. L. Goeckel and D. Towsley (all with the Univ. of Massachusetts, Amherst, USA), N. Choungmo Fofack proposes in [81] approximate models to assess the performance of a cache network with arbitrary topology where nodes run the Least Recently Used (LRU), First-In First-Out (FIFO), or Random (RND) replacement policies on arbitrary size caches. The authors take advantage of the notions of “cache characteristic time” and “Time-To-Live (TTL)-based cache” to develop a unified framework for approximating metrics of interest of interconnected caches. This approach is validated through event-driven simulations, and when possible, compared to the existing *a-NET* model.

6.6.3. Data placement and retrieval in distributed/peer-to-peer systems

Distributed systems using a network of peers have become an alternative solution for storing data. These systems are based on three pillars: data fragmentation and dissemination among the peers, redundancy mechanisms to cope with peers churn and repair mechanisms to recover lost or temporarily unavailable data. In previous years, A. Dandoush, S. Alouf and P. Nain have studied the performance of peer-to-peer storage systems in terms of data lifetime and availability using the traditional redundancy schemes. This work has now been published in [23].

A. Jean-Marie and O. Morad (Univ. Montpellier 2) have proposed a control-theoretic model for the optimization of prefetching in the context of hypervideo or, more generally, connected documents. The user is assumed to move randomly from document to document, and the controller attempts at downloading in advance the documents accessed. A penalty is incurred when the document is not completely present. The model is flexible in the sense that it allows several variants for the network model and the cost metric [63]. They have proposed exact algorithms and heuristics for the solution of this problem, and compared them on a benchmark of different user behaviors [62].

The question of whether it is possible to prefetch documents so that the user never experiences blocking, has been modeled with a “cops-and-robbers” game jointly with F. Fomin (Univ. Bergen), F. Giroire and N. Nisse (both from Inria project-team COATI) and D. Mazauric (former PhD student in MAESTRO and MASCOTTE) [24] (see also MAESTRO’s 2011 activity report).

6.6.4. Streaming optimization

In streaming applications such as youtube, packets have to be played at the destination at the same rate they were created. If a packet is not available at the destination when it has to be played then a starvation occurs. This results in an unpleasant frozen screen and in an interruption in the video. To decrease the probability of a starvation the destination first waits till it has received some target number of packets and only then starts to play them. In [32], E. Altman and M. Haddad together with Y. Xu (Fudan Univ. China), R. El-Azouzi and T. Jimenez (Univ. of Avignon), and S.-E. Elayoubi (Orange Labs, Issy les Moulineaux) compute the starvation probability as a function of the initial buffering and study tradeoffs between the two performance measures: starvation probabilities and the pre-buffering delay.

6.6.5. Stochastic geometry and network coding for distributed storage

In [37] E. Altman and K. Avrachenkov in collaboration with J. Goseling (Twente Univ., The Netherlands) consider storage devices located in the plane according to a general point process and specialize the results for the homogeneous Poisson process. A large data file is stored at the storage devices, which have limited storage capabilities. Hence, they can only store parts of the data. Clients can contact the storage devices to retrieve the data. The expected costs of obtaining the complete data under uncoded or coded data allocation strategies are compared. It is shown that for the general class of cost measures where the cost of retrieving data is increasing with the distance between client and storage devices, coded allocation outperforms uncoded allocation. The improvement offered by coding is quantified for two more specific classes of performance measures. Finally, the results are validated by computing the costs of the allocation strategies for the case that storage devices coincide with currently deployed mobile base stations.

6.7. Advances in Methodological Tools

Participants: Eitan Altman, Konstantin Avrachenkov, Ilaria Brunetti, Ioannis Dimitriou, Mahmoud El Chamie, Majed Haddad, Alain Jean-Marie, Philippe Nain, Giovanni Neglia.

6.7.1. Queueing theory

In [21] K. Avrachenkov and P. Nain in collaboration with U. Yechiali (Tel Aviv Univ., Israel) study a retrial queueing system with two independent Poisson streams of jobs flowing into a single-server service system, having a limited common buffer that can hold at most one job. If a type-i job ($i = 1, 2$) finds the server busy, it is blocked and routed to a separate type-i retrial (orbit) queue that attempts to re-dispatch its jobs at its specific Poisson rate. This creates a system with three dependent queues. Such a queueing system serves as a model for two competing job streams in a carrier sensing multiple access system. They study the queueing system using multi-dimensional probability generating functions, and derive its necessary and sufficient stability conditions while solving a Riemann-Hilbert boundary value problem. Various performance measures are calculated and numerical results are presented. In particular, numerical results demonstrate that the proposed multiple access system with two types of jobs and constant retrial rates provides incentives for the users to respect their contracts.

In [19] K. Avrachenkov in collaboration with E. Morozov (Petrozavodsk State Univ., Russia) consider a finite buffer capacity GI/GI/c/K-type retrial queueing system with constant retrial rate. The system consists of a primary queue and an orbit queue. The primary queue has c identical servers and can accommodate up to K jobs (including c jobs under service). If a newly arriving job finds the primary queue to be full, it joins the orbit queue. The original primary jobs arrive to the system according to a renewal process. The jobs have i.i.d. service times. The head of line job in the orbit queue retries to enter the primary queue after an exponentially distributed time independent of the length of the orbit queue. Telephone exchange systems, medium access protocols, optical networks with near-zero buffering and TCP short-file transfers are some telecommunication applications of the proposed queueing system. The model is also applicable in logistics. They establish sufficient stability conditions for this system. In addition to the known cases, the proposed model covers a number of new particular cases with the closed-form stability conditions. The stability conditions that they obtained have clear probabilistic interpretation.

In [20] K. Avrachenkov in collaboration with E. Morozov and R. Nekrasova (Petrozavodsk State Univ., Russia) and B. Steyaert (Ghent Univ., Belgium) study a retrial queueing system with N classes of customers, where a class-i blocked customer joins orbit i. Orbit i works like a single-server queueing system with (exponential) constant retrial time (with rate μ_{0i}) regardless of the orbit size. Such a system is motivated by multiple telecommunication applications, for instance wireless multi-access systems, and transmission control protocols. First, they present a review of some corresponding recent results related to a single-orbit retrial system. Then, using a regenerative approach, they deduce a set of necessary stability conditions for such a system. They will show that these conditions have a very clear probabilistic interpretation. They also performed a number of simulations to show that the obtained conditions delimit the stability domain with a remarkable accuracy, being in fact the (necessary and sufficient) stability criteria, at the very least for the 2-orbit M/M/1/1-type and M/Pareto/1/1-type retrial systems that they focus on.

In [75], I. Dimitriou investigates a single server system accepting two types of retrial customers and paired services. The service station can handle at most one customer, and if upon arrival a customer finds the server busy it is routed to an infinite capacity orbit queue according to its type. Upon a service completion epoch, if at least one orbit queue is non-empty, the server seeks to find customers from the orbits. If both orbit queues are non-empty, the seeking process will bring to the service area a pair of customers, one from each orbit. If only one is non-empty, then a customer from this orbit queue will be brought to the service area. However, if a primary customer arrives during the seeking process it will occupy the server immediately. It is shown that the joint stationary orbit queue length distribution at service completion epochs, can be determined via transformation to a Riemann boundary value problem. Stability condition is investigated, while an extension of the model is also discussed and analyzed. Numerical results are obtained and yield insight into the behavior of the system. The theoretical system can be used to model a relay node for two connections in wireless communication, where network coding is used.

When individuals have to take a decision on whether or not to join a queue, one may expect to have threshold equilibria in which customers join the queue if its size is smaller than a threshold and do not join if it exceeds the threshold. In [74], P. Wiecek (Wroclaw Univ. of Technology, Poland), E. Altman and A. Ghosh (Univ. of Pennsylvania, USA) have studied queueing in which the congestion cost per user decreases in the queue size. An example for such a situation is multicast communication where all individuals that participate in the multicast session share the transmission cost. They showed that many equilibria exist and computed the asymptotic system behavior as the arrival rate of individuals grows.

6.7.2. Markov processes

In [16] K. Avrachenkov in collaboration with A. Eshragh (Univ. of Adelaide, Australia) and J. Filar (Flinders Univ., Australia) present some algebraic properties of a particular class of probability transition matrices, namely, Hamiltonian transition matrices. Each matrix P in this class corresponds to a Hamiltonian cycle in a given graph G on n nodes and to an irreducible, periodic, Markov chain. They show that a number of important matrices traditionally associated with Markov chains, namely, the stationary, fundamental, deviation and the hitting time matrix all have elegant expansions in the first $n-1$ powers of P, whose coefficients can be explicitly derived. They also consider the resolvent-like matrices associated with any given Hamiltonian cycle and its reverse cycle and prove an identity about the product of these matrices. As an illustration of these analytical results, they exploit them to develop a new heuristic algorithm to determine a non-Hamiltonicity of a given graph.

6.7.3. Control theory

In [17] K. Avrachenkov and O. Habachi (former post-doc in MAESTRO) in collaboration with A. Piunovskiy and Y. Zhang (both from the Univ. of Liverpool, UK) investigate infinite-horizon deterministic optimal control problems with both gradual and impulsive controls, where any finitely many impulses are allowed simultaneously. Both discounted and long-run time-average criteria are considered. They establish very general and at the same time natural conditions, under which the dynamic programming approach results in an optimal feedback policy. The established theoretical results are applied to the Internet congestion control, and by solving analytically and non-trivially the underlying optimal control problems, they obtain a simple threshold-based active queue management scheme, which takes into account the main parameters of the transmission control protocols, and improves the fairness among the connections in a given network.

6.7.4. Game theory

6.7.4.1. Estimating the Shapley-Shubik index

In [15] K. Avrachenkov in collaboration with L. Cottatellucci (EURECOM) and L. Maggi (CREATE-NET, Italy) consider simple Markovian games, in which several states succeed each other over time, following an exogenous discrete-time Markov chain. In each state, a different simple static game is played by the same set of players. They investigate the approximation of the Shapley-Shubik power index in simple Markovian games (SSM). They prove that an exponential number of queries on coalition values is necessary for any deterministic algorithm even to approximate SSM with polynomial accuracy. Motivated by this, they propose and study three randomized approaches to compute a confidence interval for SSM. They rest upon two different assumptions, static and dynamic, about the process through which the estimator agent learns the coalition values. Such approaches can also be utilized to compute confidence intervals for the Shapley value in any Markovian game. The proposed methods require a number of queries, which is polynomial in the number of players in order to achieve a polynomial accuracy.

6.7.4.2. Evolutionary games

Evolutionary games attempt to explain the evolution of species and the dynamics of competition. The player's utility is called "fitness" and a larger fitness indicates a larger rate of reproducibility. In standard evolutionary games, one studies interactions between individuals each of which is consider as a player. In [49], I. Brunetti, E. Altman, and R. El-Azouzi (Univ. of Avignon) argue that in many situations both in biology as well as in networking, one cannot attribute a fitness to an individual but rather to a group of individuals that behaves as an altruistic entity. For example, in a hive of bees it is only the queen that reproduces and thus one cannot model a single bee as a selfish player. They present new definitions for evolutionary games for such situations and study their equilibrium.

This, as well as other considerations in multi-population evolutionary games, is applied in [56] by H. Gaiech and R. El-Azouzi (Univ. of Avignon), M. Haddad, E. Altman and I. Mabrouki (Univ. of Manouba, Tunisia) to Multiple Access Control for which the equilibrium is explicitly computed.

In [84] E. Altman presents a summary of the foundations of classical evolutionary games addressed to a wide public. Both the equilibrium notion of ESS (Evolutionary Stable Strategy) as well as the replicator dynamics (which describes the non-equilibrium behavior) are presented.

6.7.4.3. Sequential Anonymous Games

Stationary anonymous sequential games are a special class of games that combines features from both population games (infinitely many players) with stochastic games. It allows studying competition in complex systems where each individual belongs to a community (which we call individual state) which may change in time as a result of actions taken by the individual. Unlike standard evolutionary games, a player does not just optimize its immediate reward (fitness) but some long term reward over the time. P. Wiecek (Wroclaw Univ. of Technology, Poland) and E. Altman proved in [42] the existence of an equilibrium for the general model and studied the two applications. The first one is described in §6.5.1 .

The second application is a maintenance repair problem: each of a large number of cars can decide whether to behave gently or to drive fast. By driving fast it takes larger risks for having an accident. The probability of an accident depends on the fraction of drivers that drive fast. An internal state of the car is either good (g) or bad (b). A car gets to a state b as a result of an accident and then it has some penalty and costs for repair. The advantage of driving fast is reducing delay costs. This problem is formulated as a sequential anonymous game and its equilibrium is computed. They computation makes use of the linear structure of both the transition probabilities and the immediate fitness in the global state.

6.7.5. Optimization

In [55] M. El Chamie and G. Neglia provide a methodology for solving smooth norm optimization problems under some linear constraints using the Newton's method. This problem arises in many machine learning and graph optimization applications. They show how Newton's method significantly outperforms gradient methods both in terms of convergence speed and in term of robustness to the step size selection.

MAGIQUE-3D Project-Team

6. New Results

6.1. Inverse Problems

6.1.1. Complex-frequency domain Full Waveform Inversion

Participants: Florian Faucher, Maarten V. de Hoop, Henri Calandra.

We study the seismic inverse problem for the (complex) frequency-domain elastic isotropic wave equation; in particular the recovery of the Lamé parameters and density. We employ a Full Waveform Inversion where the iterative minimization is based on a gradient descent. The elastic inverse problem shows a Lipschitz-type stability where the Fréchet derivative has a strictly positive 'lower bound'. This bound is connected to the stability constant and can be approximated using the Gauss-Newton Hessian. The successive stability estimates provide a control of the convergence and decide the parameters of inversion. We develop a multi-level approach based on a structured domain partitioning of the sub-surface. The coefficients (Lamé parameters and density) are assumed to be piecewise constant functions following the domain partitioning, which is naturally defined with the successive stability estimates to maintain the radius of convergence, while refinement provides resolution. It allows us to start with minimal prior information for the coefficients and the algorithm is perfectly suitable for complex frequency. We have carried out numerical experiments in two and three dimensions; those results have been presented during the following conferences in 2014: [48], [49].

6.1.2. Imaging of complex media with elastic wave equations

Participants: Jérôme Luquel, Hélène Barucq, Henri Calandra, Julien Diaz.

Even if RTM has enjoyed the tremendous progresses of scientific computing, its performances can still be improved, in particular when applied to strong heterogeneous media. In this case, images have been mainly obtained by using direct arrivals of acoustic waves and the transition to elastic waves including multiples is not obvious essentially because elastic waves equations are still more computationally consuming. The accuracy of numerical wave fields is obviously of great importance. We have thus chosen to consider high-order Discontinuous Galerkin Methods which are known to be well-adapted to provide accurate solutions based upon parallel computing. Now one of the main drawback of RTM is the need of storing a huge quantity of information which is redhibitory when using elastic waves. For that purpose, we apply the Griewank algorithm following Symes' ideas for the acoustic RTM. The idea is to find a compromise between the number of wave equations to solve and the number of numerical waves that we have to store. This is the so-called Optimal Checkpointing. By reducing the occupancy of the memory, RTM should be efficient even when using elastic waves. By this way, one may wonder if considering elastic waves including multiples in order to improve images of heterogeneous media is a valid option. It must involve a careful numerical analysis including the evaluation of the impact of the imaging condition. It is thus necessary to derive accurate imaging conditions, which could take advantage of all the information contained in the wavefield. For acoustic media, Claerbout proposed an imaging condition which is widely used and turns out to be sufficient to accurately reproduce interfaces. But Claerbout conditions do not take wave conversions into account and it is not clear wether conversions do or do not contain interesting information to get accurate images of heterogeneous media.

Since P-wave and S-wave interact with each other, it might be relevant to use an imaging condition including these interactions. In fact, this has been done successfully by J.Tromp and C. Morency for seismology applications based upon the inversion of the global Earth. Their approach is based upon the state adjoint and it involves sensitivity kernels which are defined from the propagated and the back-propagated fields. Now it has been shown that full wave form inversions using these sensitivity kernels may be polluted by numerical artefacts. One solution is to use a linear combination of the sensitivity kernels to delete artefacts. In this work, we propose then a new imaging condition which construction is inspired from with some approximations required to keep admissible computational costs. We illustrate the properties of the new imaging condition on industrial benchmarks like the Marmousi model. In particular, we compare the new imaging condition with other imaging conditions by using as criteria the quality of the image and the computational costs required by the RTM. This work was presented at the the WCCM XI - ECCM V - ECFD VI - Barcelona 2014 Conference and SIAM Conference on IMAGING SCIENCE (SIAM-IS14) Hong Kong Baptist University[67].

6.1.3. Helioseismology

Participants: Juliette Chabassier, Marc Duruflé, Thorsten Hohage.

We have begun to write a software interface that allows to solve an inverse problem using adjoint and regularization methods (iTReg software) while using Montjoie software for the direct problem that must be solved at each iteration of the inversion process.

6.2. Modeling

6.2.1. High-Order Time Schemes

6.2.1.1. Fourth order energy-preserving locally implicit discretization for linear wave equations

Participants: Juliette Chabassier, Sébastien Imperiale.

A family of fourth order coupled implicit-explicit schemes is presented as a special case of fourth order coupled implicit schemes for linear wave equations. The domain of interest is decomposed into several regions where different fourth order time discretization are used, chosen among a family of implicit or explicit fourth order schemes derived in [72]. The coupling is based on a Lagrangian formulation on the boundaries between the several non conforming meshes of the regions. A global discrete energy is shown to be preserved and leads to global fourth order consistency. Numerical results in 1d and 2d illustrate the good behavior of the schemes and their potential for the simulation of realistic highly heterogeneous media or strongly refined geometries, for which using everywhere an explicit scheme can be extremely penalizing. Accuracy up to fourth order reduces the numerical dispersion inherent to implicit methods used with a large time step, and makes this family of schemes attractive compared to second order accurate methods in time. This work has been presented at the Franco-Russian workshop on mathematical geophysics, Sep 2014, Novosibirsk, Russia [58], at the and is the object of a submitted publication to International Journal for Numerical Methods in Engineering.

6.2.1.2. A new modified equation approach for solving the wave equation

Participants: Hélène Barucq, Henri Calandra, Julien Diaz, Florent Ventimiglia.

In order to obtain high-order time-schemes, we are considering an alternative approach to the ADER schemes and to the modified equation technique described in section 3.2 . The two first steps of the construction of the schemes are similar to the previous schemes : we apply a Taylor expansion in time to the solution of the wave equation and we replace the high-order derivatives with respect to the time by high order space operators, using the wave equation. The difference is that we do not use auxiliary variables and we choose to discretize directly the high-order operators in space.

In the framework of the PhD thesis of Florent Ventimiglia, we have extended this new method involving p-harmonic operator to the first order formulation of the acoustic wave equation, which is the formulation discretized in the DIVA platform of TOTAL. In this case, the high order operators in space are not are not powers of the Laplace operator but powers of the gradient. Hence, we also had to adapt the space discretization, and we have extended the DG formulation with centered fluxes proposed in [77] to higher order operators. A numerical analysis of performance in 2D indicates that, for a given accuracy, this method requires less computational costs and less storage than the High-Order ADER Scheme. These results have been presented to the AIMS conference [54]. A paper has been published in ESAIM Proceedings [19].

6.2.2. Finite Element Methods for the time-harmonic wave equation.

6.2.2.1. Goal-Oriented Adaptivity using Unconventional Error Representations

Participants: Vincent Darrigrand, David Pardo, Ignacio Muga, Hélène Barucq.

In the scope of subsurface modelling via the resolution of inverse problems, the so-called goal-oriented adaptivity plays a fundamental role. Indeed, while classical adaptive algorithms were first designed to accurately approximate the energy norm of a problem [69], [70], one requires a good approximation of a specific quantity of interest. An energy norm driven self-adaptive strategy can still be used for that purpose, although it often becomes sub-optimal and unable to provide an accurate solution for the required quantity of interest in a reasonable amount of time.

During the late 90's, to overcome this issue, the so-called goal-oriented strategy appeared, see for instance [82], [81]. The goal-oriented approach consists in expressing the error in the quantity of interest as an integral over the entire computational domain involving the errors of the original and adjoint problems, and then minimise an upper bound of such error representation by performing local refinements.

Most authors, using the adjoint problem, represent the approximation error in the quantity of interest via the global bilinear form that describes the problem in terms of local and computable quantities.

Our methodology, however, is based on the selection of an alternative bilinear form exhibiting better properties than the original bilinear form (e.g. positive definiteness). We represent the residual error functional of the adjoint problem through this alternative form. We can then compute new upper bounds of the error of the quantity of interest in a similar way than with the classical approach. Our main goal is to demonstrate that a proper choice of such alternative form may improve the upper bounds of the error representation.

Moreover, the method proposed here generalises the existing ones, since, in particular, we can select as the alternative bilinear form the one associated to the adjoint problem.

6.2.2.2. *Hybridizable Discontinuous Galerkin method for the elastic Helmholtz equations*

Participants: Marie Bonnasse-Gahot, Henri Calandra, Julien Diaz, Stéphane Lanteri.

We consider Discontinuous Galerkin (DG) methods formulated on fully unstructured meshes, which are more convenient than finite difference methods on cartesian grids to handle the topography of the subsurface. DG methods and classical Finite Element (FE) methods mainly differ from discrete functions which are only piecewise continuous in the case of DG approximation. DG methods are then more suitable than Continuous Galerkin (CG) methods to deal with hp-adaptivity. This is a great advantage to DG method which is thus fully adapted to calculations in highly heterogeneous media. Nevertheless, the main drawback of classical DG methods is that they are more expensive in terms of number of unknowns than classical CG methods, especially when arbitrarily high order interpolation of the field components is used. In this case DG methods lead to larger sparse linear systems with a higher number of globally coupled degrees of freedom as compared to CG methods with a same given mesh. In that case, we consider a hybridizable Discontinuous Galerkin (HDG) method which principle consists in introducing a Lagrange multiplier representing the trace of the numerical solution on each face of the mesh cells. This new variable exists only on the faces of the mesh and the unknowns of the problem depend on it. This allows us to reduce the number of unknowns of the global linear system. Now the size of the matrix to be inverted only depends on the number of the faces of the mesh and on the number of the degrees of freedom of each face. It is worth noting that for the classical DG method it depends on the number of the cells of the mesh and on the number of the degrees of freedom of each cell. The solution to the initial problem is then recovered thanks to independent elementwise calculation. The principle of the HDG method and 2D results were presented at the WCCM XI - ECCM V - ECFD VI - Barcelona 2014 Conference [41], the EAGE Workshop on High Performance Computing for Upstream [42], the Second Russian-French Workshop "Computational Geophysics" [43] and at the Réunion des Sciences de la Terre 2014 conference [53]. A comparison between HDG method and classical nodal DG method was given on a poster at the Journées Total-Mathias 2014 workshop [66].

6.2.2.3. *Helioseismology*

Participants: Hélène Barucq, Juliette Chabassier, Marc Duruflé, Damien Fournier, Laurent Gizon.

The finite element code Montjoie 5.2 has been used to solve Helmholtz equation in axisymmetric domain in the configuration of the sun. The efficiency of the code has been compared in three configurations : radial (1-D mesh and spherical harmonics), axisymmetric (2-D mesh), 3-D. The results have convinced our-selves and our partners of Max Planck Institute that the axisymmetric configuration is the most interesting for an inversion procedure, since 3-D computations are too expensive. A more realistic modeling of the sun requires the solution of time-harmonic Galbrun's equations (instead of Helmholtz equation), different formulations have been implemented and studied. It appeared that the different numerical methods are not able to converge to the correct solution for non-uniform flows. The lack of convergence is more obvious for flows with a larger Mach number. Such problems do not appear in Linearized Euler equations, as a result we have proposed simplified Galbrun's equations that converge correctly and provide the same solution as original Galbrun's equations for a null flow. These equations have been implemented in 2-D, axisymmetric and 3-D configuration.

6.2.2.4. Scattering of acoustic waves by a disc - Hypersingular integral equations

Participants: Leandro Farina, Paul Martin, Victor Péron.

Two-dimensional boundary-value problems involving a Neumann-type boundary condition on a thin plate or crack can often be reduced to one-dimensional hypersingular integral equations. Examples are potential flow past a rigid plate, acoustic scattering by a hard strip, water-wave interaction with thin impermeable barriers, and stress fields around cracks. In [29], we generalize some of these results to two-dimensional hypersingular integral equations. Thus, rather than integrating over a finite interval, we now integrate over a circular disc. Two-dimensional hypersingular equations over a disc arise, for example, in the scattering of acoustic waves by a hard disc; this particular application is described in Appendix A. We develop an appropriate spectral (Galerkin) method, using Fourier expansions in the azimuthal direction and Jacobi polynomials in the radial direction. The Hilbert-space arguments used by Golberg are generalized and a convergence theorem is proved by using tensor-product techniques. Our results are proved in weighted L^2 spaces. Then, Tranter's method is discussed. This method was devised in the 1950s to solve certain pairs of dual integral equations. It is shown that this method is also convergent because it leads to the same algebraic system as the spectral method.

6.2.2.5. Finite Element Subproblem Method

Participants: Patrick Dular, Christophe Geuzaine, Laurent Krähenbühl, Victor Péron.

In the paper [26], the modeling of eddy currents in conductors is split into a sequence of progressive finite element subproblems. The source fields generated by the inductors alone are calculated at first via either the Biot-Savart law or finite elements. The associated reaction fields for each added conductive region, and in return for the source regions themselves when massive, are then calculated with finite element models, possibly with initial perfect conductor and/or impedance boundary conditions to be further corrected. The resulting subproblem method allows efficient solving of parameterized analyses thanks to a proper mesh for each subproblem and the reuse of previous solutions to be locally corrected.

6.2.2.6. High Order Methods for Helmholtz Problems in Highly Heterogeneous Media

Participants: Théophile Chaumont-Frelet, Henri Calandra, Hélène Barucq, Christian Gout.

Heterogeneous Helmholtz problems arise in various geophysical application where they modelize the propagation of time harmonic waves through the subsurface. For example, in inversion problems, the aim is to reconstruct a map of the underground based on surface acquisition. This recovery process involves the solution to several Helmholtz problems set in different media, and high frequency solutions are required to obtain a detailed image of the underground. This observations motivate the design of efficient solver for highly heterogeneous Helmholtz problems at high frequency.

The main issue with the discretization of high frequency problems is the so called "pollution effect" which impose drastic condition on the mesh. In the homogeneous case, it is known that one efficient way to reduce the pollution effect is the use of high order discretization methods. However, high order methods can not be applied as is to highly heterogeneous media. Indeed, they are based on coarser mesh and are not sensitive to fine scale variations of the medium.

We propose to overcome this difficulty by using a multiscale strategy to take into account fine scale heterogeneities on coarse meshes. The method is based on a simple medium approximation method, which can be seen as a special quadrature rule. Numerical experiments in two dimensional geophyscial benchmarks show that high order method coupled with our multiscale approximation medium stragey are cheaper than low order method for a given accuracy. Futhermore, focusing on one dimensional models, we were able to show from a theoretical point of view that our methology reduces the pollution effect even when used on coarse meshes with non-matching interfaces.

This work has been presented at the *WCCM XI - ECCM V - ECFD VI - Barcelona 2014* conference, the *Second Russian-French Workshop "Computational Geophysics"*. A poster has been presented at the *journées Total-Mathias 2014 workshop*. A paper has been submitted for publication to *Math. Of Comp.*.

6.2.3. Boundary conditions.

6.2.3.1. Absorbing Boundary Conditions for Tilted Transverse Isotropic Elastic Media

Participants: Lionel Boillot, Hélène Barucq, Julien Diaz, Henri Calandra.

The seismic imaging simulations are always performed in bounded domains whose external boundary does not have physical meaning. We have thus to couple the wave equations with boundary conditions which aim at reproduce the invisibility of the external boundary. The discretization of these conditions can be an issue. For instance, an efficient condition, once discretized, can induce huge computational costs by filling the matrix which has to be inverted. This is the case of the transparent boundary conditions which are approximated by local Absorbing Boundary Conditions (ABC) that do not increase to much the computational burden. However, the ABC has the drawback to introduce spurious numerical waves which can perturb the RTM results. It is possible to avoid this drawback by applying PML (Perfectly Matched Layers) but it proves to be unstable in anisotropic media. Last year, we proposed a way of construction leading to a stable ABC. The technique is based on slowness curve properties, giving to our approach an original side. We established stability results from long time energy behavior and we have illustrated the performance of the new condition in 2D numerical tests. This year, we extend all these results to 3D case and to arbitrary boundary shapes. The previous paper submission on 2D results has been accepted and released [18]. The recent results in 3D have been presented to the ECCOMAS conference.

6.2.3.2. Derivation of high order absorbing boundary conditions for the Helmholtz equation in 2D.

Participants: Hélène Barucq, Morgane Bergot, Juliette Chabassier, Élodie Estecahandy.

Numerical simulation of wave propagation raises the issue of dealing with outgoing waves. In most of the applications, the physical domain is unbounded and an artificial truncation needs indeed to be carried out for applying numerical methods like finite element approximations. Adapted boundary conditions that avoid the reflection of outgoing waves and provide a well-posed mathematical problem must then be derived. With ideal boundary conditions, the solution on the new mixed boundary valued problem in the truncated domain would actually be equal to the restriction of the mathematical solution in the unbounded domain. However, such ideal boundary conditions, called "transparent boundary conditions", can be shown to be nonlocal, which leads to dramatic computational overcosts. The seek of local boundary conditions, called "absorbing boundary conditions" (ABC), has been the object of numerous works trying to perform efficient conditions based on different techniques of derivation. Among them, the technique of micro-diagonalisation has been employed to the wave equation and more generally to hyperbolic systems in [76], leading to a hierarchy of absorbing local boundary conditions based on the approximation of the Dirichlet-to-Neumann map. A comprehensive review of different used strategies and higher order conditions can be found in [85]. One desirable property of ABCs is that the reflection of the waves on the artificial boundary generates an error of the same order as the one generated by the spatial discretization inside the domain. The computational effort is thus optimized in terms of modeling and numerical inaccuracies. Moreover, the ABC must fit the artificial boundary chosen by the user of the method. In the context of high order spatial discretization (spectral finite elements [74], Interior Penalized Discontinuous Galerkin [68]), there is nowadays a need for high order ABCs that can adapt on non flat geometries since these methods prove very efficient for capturing arbitrary shaped domains.

The aim of the present work is to develop high order ABCs for the Helmholtz equation, that can adapt to regular shaped surfaces. A classical way of designing ABCs is to use Nirenberg theorem [80] on the second order formulation of the Helmholtz equation, which enables us to decompose the operator as a product of two first order operators. Here our approach is to rewrite the Helmholtz equation as a first order system of equations before developing ABCs using M.E. Taylor's micro-diagonalisation method [84]. Then an asymptotic truncation must be performed in order to make the ABC local, and we will see that the high frequency approximation will lead to more usable ABCs than the one stating that the angle of incidence is small. During the process, while increasing the degree of the pseudo differential operator decomposition along with the order of asymptotic truncation, we retrieve classical ABCs that have been found with other techniques by other authors. For now, we have restricted ourselves to two dimensions of space, but despite the fact that 3D generalization should obviously generate more calculation, no further theoretical difficulties are expected.

This work has been the object of a technical report [61] and the obtained conditions have been implemented in Montjoie 5.2 and Houd10ni 5.1 .

6.2.4. Asymptotic modeling.

6.2.4.1. Fast Simulation of Through-casing Resistivity Measurements Using Semi-analytical Asymptotic Models.

Participants: Victor Péron, David Pardo, Aralar Erdozain.

When trying to obtain a better characterization of the Earth's subsurface, it is common to use borehole through-casing resistivity measurements. It is also common for the wells to be surrounded by a metal casing to protect the well and avoid possible collapses. The presence of this metal case highly complicates the numeric simulation of the problem due to the high conductivity of the casing compared to the conductivity of the rock formations. In this study [47] we present an application of some theoretical asymptotic methods in order to deal with complex borehole scenarios like cased wells. The main idea consists in replacing the part of the domain related to the casing by a transmission impedance condition. The small thickness of the casing makes it ideal to apply this kind of mathematical technique. When eliminating the casing from the computational domain, the computational cost of the problems considerably decreases, while the effect of the casing does not disappear due to the impedance transmission conditions. The results show that when applying an order three impedance boundary condition for a simplified domain, it only generates a negligible approximation error, while it considerably reduces the computational cost. For obtaining the numerical results and testing the mathematical models we have developed a Finite Element Code in Matlab. The code works with Lagrange polynomials of any degree as basis functions and triangular shaped elements in two dimensions. The code has been adapted for working with the transmission impedance conditions required by the mathematical models.

6.2.4.2. Modeling the propagation of ultrashort laser pulses in optical fibers.

Participants: Mohamed Andjar, Juliette Chabassier, Marc Duruflé.

In order to model the propagation of an ultrashort laser pulse, the most natural idea is to solve Maxwell's equations in a nonlinear and dispersive medium. Given the considered optical periods (around 10^{-14} seconds), the associated wavelengthes (around 1 millimeter) and the propagation distances (several meters), the direct numerical simulation of these equations by usual numerical techniques (finite elements, explicit time schemes) is impossible because too expensive. The standard procedure is therefore to use approached equations obtained by exploiting legitimate hypotheses in the considered context (slowly varying pulse envelope, narrow spectrum, paraxial approximation ...). These new equations, among them the Nonlinear Schrödinger Equation, are significantly less expensive to solve and we can therefore provide realistic numerical simulations to physicists.

When the pulse propagates in an optical fiber, its spatial profile in the orthogonal plane to the propagation direction in very simple because optical fibers posses a finite (small, often equal to one) number of propagating modes. The equations that originally are stated on a 3D domain can then be written as one spatial dimension equations.

The scientific objective of this internship was to apply the approximation techniques mentioned above in this specific context, in order to obtain one or several equations (depending on the used hypotheses) that model the propagation of ultrashort laser pulses in optical fibers. A matlab code has been developed and integrated in the C++ code Montjoie 5.2 . Numerical simulations have been led in order to observe classical situations of nonlinear fiber optics (Kerr effect, Raman effect, supercontinuum generation, ...).

6.2.4.3. Small heterogeneities in the context of time-domain wave propagation equation : asymptotic analysis and numerical calculation

Participants: Vanessa Mattesi, Sébastien Tordeux.

We have focused our attention on the modeling of heterogeneities which are smaller than the wavelength. The work can be decomposed into two parts : a theoretical one and a numerical one. In the theoretical one, we derive a matched asymptotic expansion composed of a far-field expansion and a near-field expansion. The terms of the far-field expansion are singular solutions of the wave equation whereas the terms of the near-field expansion satisfy quasistatic problems. These expansions are matched in an intermediate region. We justify mathematically this theory by proving error estimates. In the numerical part, we describe the Discontinuous Galerkin method, a local time stepping method and the implementation of the matched asymptotic method. Numerical simulations illustrate these results. Vanessa Mattesi has defended her PhD on this topic[14].

6.2.4.4. Theoretical and numerical investigations of acoustic response of a multiperforated plate for combustion liners

Participants: Vincent Popie, Estelle Piot, Sébastien Tordeux.

Multiperforated plates are used in combustion chambers for film cooling purpose. As the knowledge of the acoustic response of the chamber is essential for preventing combustion instabilities, the acoustic behaviour of the perforated plates has to be modeled. This can be done either by considering the transmission impedance of the plates, or their Rayleigh conductivity.

We have investigated the link between these two quantities thanks to matched asymptotic expansions. Especially the far-field or near-field nature of the physical quantities used in the definition of the impedance and Rayleigh quantity has been enlightened. Direct numerical simulations of the propagation of an acoustic plane wave through a perforated plate are performed and post-treated so that the assumptions underlying the definitions of impedance and Rayleigh conductivity have been checked. The results will be presented at the conference ASME Turbo Expo 2015.

6.3. High Performance methods for solving wave equations

6.3.1. Coupling the DG code with task programming libraries

Participants: Lionel Boillot, Emmanuel Agullo, George Bosilca, Henri Calandra.

The parallelization of the original code is based on a preliminary step of domain decomposition and then on the use of the MPI (Message Passing Interface) library. It is a common choice which works pretty well in most of the classical architectures. However, the parallel efficiency is not optimal and the performance decreases in hybrid architectures. Indeed, we know the number of operations that each sub-domain has to performed but this does not give us the exact time that the computations require. The cluster heterogeneity leads to various automatic optimizations (memory cache, parallel capability, . . .) which are difficult to measure. We have decided to tackle this problem by modifying the parallelism with the use of task programming. We have thus rewritten the DIVA algorithm in a graph of tasks without using the MPI library and we have left to the runtime PaRSEC the choice of when and where to execute each task. The numerical experiments we have performed have confirmed the significant improvement of the parallel efficiency on different architectures like ccNUMA machines or Intel Xeon Phi co-processors. Moreover, the proposed solution is portable on these architectures, this means that none or few modifications are required in the code, allowing to focus on algorithmic aspects in order to preserve the performance. These results have been presented to the EAGE HPC workshop and to the HPCC IEEE conference within a paper have been accepted.

MAGNET Team

6. New Results

6.1. Highlights of the Year

We developed a new framework for high order learning [4].

We have illustrated the usefulness of automatically annotated examples in complex learning supervised by few training examples [2], [1].

We propose a new algorithm for semi-supervised spectral clustering and apply it to the NLP task of noun phrase coreference resolution [6].

6.2. Higher-order Learning with Graphs

Along the thesis of THOMAS RICATTE, in [4] and [8], we propose methods for learning from interactions between groups in networks. We propose a proper extension of graphs, called hypernode graphs as a formal tool able to model group interactions. A hypernode graph is a collection of weighted relations between two groups of distinct nodes. Weights quantify the individual participation of nodes to a given relation. We define Laplacians and kernels for hypernode graphs and and prove that they strictly generalize over graph kernels and hypergraph kernels. We prove that hypernode graphs correspond to signed graphs such that the matrix $D - W$ is positive semidefinite. As a consequence, homophilic relations between groups may lead to non homophilic relations between individuals. We define the notion of connected hypernode graphs and a resistance distance for connected hypernode graphs. We propose spectral learning algorithms on hypernode graphs allowing to infer node ratings or node labelings. As a proof of concept, we model multiple players games with hypernode graphs and we define skill rating algorithms competitive with specialized algorithms.

6.3. Natural Language Processing

In [6] (presented by DAVID CHATEL at the ECML–PKDD and CAp'2014 conferences) we propose a new algorithm for semi-supervised spectral clustering and apply it to the task of noun phrase coreference resolution. The main insight is in the inclusion of pairwise constraints into spectral clustering: our algorithm learns a new representation space for the data together with a distance in this new space. The representation space is obtained through a constraint-driven linear transformation of a spectral embedding of the data, and constraints are expressed with a Gaussian function that locally reweights the similarities in the projected space. A global, non-convex optimization objective is then derived and the model is learned via gradient descent techniques. Our algorithm is evaluated on the CoNLL-2012 coreference resolution shared task dataset, and shows some encouraging results.

In [2] and [1], we develop a new approach for the automatic identification of so-called implicit discourse relations. Specifically, our system combines hand labeled examples and automatically annotated examples based on explicit relations using several simple methods inspired by work in domain adaptation. Our system is evaluated empirically on the Annodis corpus, a French corpus annotated with discourse structures. Our system yields significant performance gains compared to only using hand-labeled data or using only automatically annotated data.

6.4. Ongoing work

6.4.1. Adaptive Graph Construction

We worked on developing a new algorithm in order to construct a graph in a adaptive way for a specific task. More precisely, we looked for a metric learning algorithm that could depend on the target task. Previous works on metric learning ([12]) aim at learning a relevant metric using a linear approach, which cannot capture the non-linearity of the data. Our approach, instead, aims at learning the most appropriate non-linear data projection for the target task. For this purpose, we train a neural network with relative constraints depending on the target task and a target classic metric (e.g. euclidean distance, cosine similarity, ...), in order to make the metric meaningful for the new data representation and our target task.

6.4.2. Correlation Clustering and Similarity/Dissimilarity Links

From a mathematical point of view, signed networks are graphs whose edges carry a sign representing the positive or negative nature of the relationship between the incident nodes. These structures are extremely useful for modeling, at the same time, similarity and dissimilarity object relationships. Given an undirected signed graphs, in the Correlation Clustering problem the goal is to find a node partition into clusters minimizing the number of negative (dissimilarity) edges linking two nodes within the same cluster and the number of positive (similarity) edges between different clusters.

We focused on devising an algorithm able to solve the Correlation Clustering problem for general input signed graphs (if the input is a complete signed graph the problem is proven to be much easier). One of the main objective of this work is the use of the proposed algorithm for creating a learner able to predict the unknown edges signs of a given signed graph. This prediction task is known as Link Classification in signed graphs. In fact, given an undirected signed graph whose edge set is split into training and test set, we could use the Correlation Clustering solution working for general input graphs for partitioning the training set and using the node partition generated for predicting the test edge signs. Moreover, one could exploit such an algorithm for developing new strategies for the Link Classification problem operating within the online and active Machine Learning setting.

Since the node set partitioning turns out to be strictly related to the Link Classification problem, we also focused on the very challenging goal of obtaining a deep understanding of the complex interplay between Link and Node Classification. More precisely, we investigated the relationships between the Vapnik Chervonenkis dimension of any given set of hypothesis space of node and edge similarity functions operating within this framework.

6.4.3. Ranking from Pairwise Sets of User Preferences

Given a set of objects (vertices of a graph) and a set of pairwise preference labels between objects (directed edges connecting vertices) which may be non-transitive due to irrationality or arbitrary noise, what is a correct way to sample preference labels for ordering the set of objects? This long standing open problem is, as far as we know, unsolved when each pairwise preference labels refers to two (disjoint) sets of objects (vertices). This framework can be easily motivated considering that quite often, in many real world contexts, users express their preferences between sets of items rather than single items, and turns out to be strictly connected with our recent model of hypergraphs with bipartite hyperedges [4]. We are working on devising a new algorithm able to rank a given set of items (graph node set) when only comparisons between sets containing at least 2 items are allowed. This challenging and interested problem is, as far as we know, quite novel and can be studied within different Machine Learning setting (online, batch, active, ...). The preliminaries results we are obtaining, when setting the cardinality of the item sets equal to 2, are encouraging and indicate that it could be possible to extend our strategies in order to deal with larger item sets.

6.5. Other results

In this section we provide the results we obtained that are not related with our main research directions.

In [3] we study the problem of learning sequential top-down tree-to-word transducers (STWs). First, we present a Myhill-Nerode characterization of the corresponding class of sequential tree-to-word transformations (STW). Next, we investigate what learning of stws means, identify fundamental obstacles, and propose a learning model with abstain. Finally, we present a polynomial learning algorithm.

MAGNOME Project-Team

6. New Results

6.1. Highlights of the Year

In collaboration with colleagues from the Institut du Vigne et du Vin (ISVV), Bordeaux and the Universidade Nova de Lisboa, Lisbon we used a population genomics approach to investigate the global phylogeography and domestication fingerprints of winemaking yeasts, using a collection of isolates obtained from fermented beverages and from natural environments on five continents. These results appeared in *Nature Communications* [11].

6.2. A Gondwanan imprint of S. uvarum diversity

Domestication of livestock and crops has been amply demonstrated through historical and archeological records, but domestication of microorganisms is much more difficult to establish. In a large-scale study [11] of the wine and cider yeast *Saccharomyces uvarum* conducted with colleagues from the Institut du Vigne et du Vin, Bordeaux and the Universidade Nova de Lisboa, Lisbon, we found the first indications of its domestication in the transition from its habitat in *Nothofagus* (southern beech) trees on the Gondwana mega-continent, to its present-day diversity in the Holarctic. The global phylogeography of these microorganisms was investigated through genome sequencing and comparison of 54 strains isolated on five continents, resulting in the identification of 10^5 high-quality SNPs and a remarkable pattern of introgressions ([11] figure 3 http://dx.doi.org/10.1038/ncomms5044).

The 54 genomes in this study were isolated, selected, and sequenced, and both assembled and aligned against reference genomes. Phylogenies were based on concatenated SNP alignment of selected chromosomes. The structure of the population was investigated using model-based Bayesian clustering.

In addition to the biological result, this study illustrates the ubiquity of an experimental approach based on large-scale sequencing of highly related genomes, in order to isolate tiny differences linked to a trait of interest. This is in contrast to the strategy that was current eight years ago, based on sequencing of a modest number of genomes spanning a much greater evolutionary range.

6.3. Improving inference of metabolic models

Participants: David James Sherman [correspondant], Pascal Durrens, Razanne Issa, Anna Zhukova.

The Pantograph approach uses reference model annotated by *gene associations*, and voting between complementary predictions of homology between reference genes and target genes, to decide whether a reaction that is present in the scaffold ought be be present in the target. A gene association implicitly represents expert knowledge about the role of genes in a compact way. If the gene association can be rewritten into a possibly satisfiable formula, then the corresponding reaction is instantiated in the target model.

Historically, gene associations have been used intuitively by experts during the model design and curation process, and are often inconsistent. We have formalized the construction of gene associations based on the semantics of different interpretations, showing how different boolean formulas should be constructed when the application is *i*) metabolic model inference, *ii*) flux-balance analysis, *iii*) hierarchical modeling, or *iv*) dynamic simulation (Razanne Issa, MS in prep.).

Second, we have refined our strategy for inferring metabolic models using abductive logic. We have shown that given a set of genes as observations in the target organism, and rules for rewriting gene associations while respecting integrity constraints for the model, then the reactions in the target model can be abduced as hypotheses that “explain” the presence of a maximial number of genes in the target genome. The advantage of this approach is that it can invent, through specialization, reactions that are not present *per se* in the reference model. Two classes of reactions can be invented: substrate-specific reactions inferred from expansion in gene families, and transport reactions needed to maintain model integrity for constitutive compartments.

6.4. Knowledge-based generalization of metabolic models

Participants: David James Sherman [correspondant], Pascal Durrens, Razanne Issa, Anna Zhukova.

Large metabolic networks are hard to understand and curate, because the large number of detailed reactions, which are needed for accurate modeling and simulation, obscure the high-level structure of the reaction network. We defined knowledge-based methods that factor similar reactions into “generic” reactions in order to visualize a whole pathway or compartment, while maintaining the underlying model so that the user can later “drill down” to the specific reactions if need be[15], [16] An implementation of this method is available as a Python library (see paragraph 5.3).

Figures 2 and 3 illustrate model generation for *Yarrowia lypolitica* fatty acid oxidation in the peroxisome. Molecular species are represented using SBGN notation: as circular nodes, and the reactions as square ones, connected by edges to their reactants and products. Ubiquitous species are of smaller size and colored gray. Non-ubiquitous species are divided into six equivalence classes, and coloured accordingly. The size of the model does not allow for readability of the species labels, thus we do not show them (figure 2).

The specific model is appropriate for simulation, because it contains all of the precise reactions. The generalized model is suited for a human, because it reveals the main properties of the model and masks distracting details. For example, the generalized model highlights the fact that there is a particularity concerning *C24:0-CoA (stearoyl-CoA)* (yellow): there exists a "shortcut" reaction (orange), producing it directly from another *fatty acyl-CoA* (yellow), avoiding the usual four-reaction beta-oxidation chain, used for other *fatty acyls-CoA*. This shortcut is not obvious in the specific model, because it is hidden among a plethora of similar-looking reactions.

We formally defined the generalization method in [15] and showed how to calculate it using a good approximation to an NP-complete set cover problem. The method was further validated in a collection of 1283 inferred models and revealed, on the one hand, a number of probable errors in the inferred models, and on the other hand, that there exist different families of generalization with a plausible link to different adaptive responses.

6.5. Characterization of STAND protein families

Participants: David James Sherman, Pascal Durrens, Witold Dyrka [correspondant].

In collaboration with Sven Saupe and Mathieu Paoletti from IBGC Bordeaux (ANR Mykimun), we worked on characterization of the STAND protein family in the fungal phylum. We established an *in silico* screen based on state-of-the-art bioinformatic tools, which – starting from experimentally studied sequences from *Podospora anserina* – allowed us to determine the first systematic picture of fungal STAND protein repertoire (ms. in preparation). Most notably, we found evidence of extensive modularity of domain associations, and signs of concerted evolution within the recognition domain [13]. Both results support the hypothesis that fungal STAND proteins, originally described in the context of vegetative incompatibility, are involved in a general fungal immune system. In addition, we investigated improved protein domain representations and elaborated a grammatical modelling method [23], which will be used to elucidate mechanisms of formation and operation of the STAND proteins.

NLR domains identified in this work have been incorporated into the upcoming release of Pfam [0].

To further explor the underlying mechanisms of repeat formation we implemented a stochastic string rewriting system that models the generation process of highly internally conserved repeats. The system is grounded in the biology of the process as it models transformation of repeats through the events of unequal crossing-over and mutation, which are believed to be main mechanisms that produce diversity in repeats. We confirmed that highly variable sites identified on the basis of entropy, are subject to selective pressure towards composition typical for binding sites, which is consistent with the suggested role of recognition epitopes.

[0] http://pfam.xfam.org

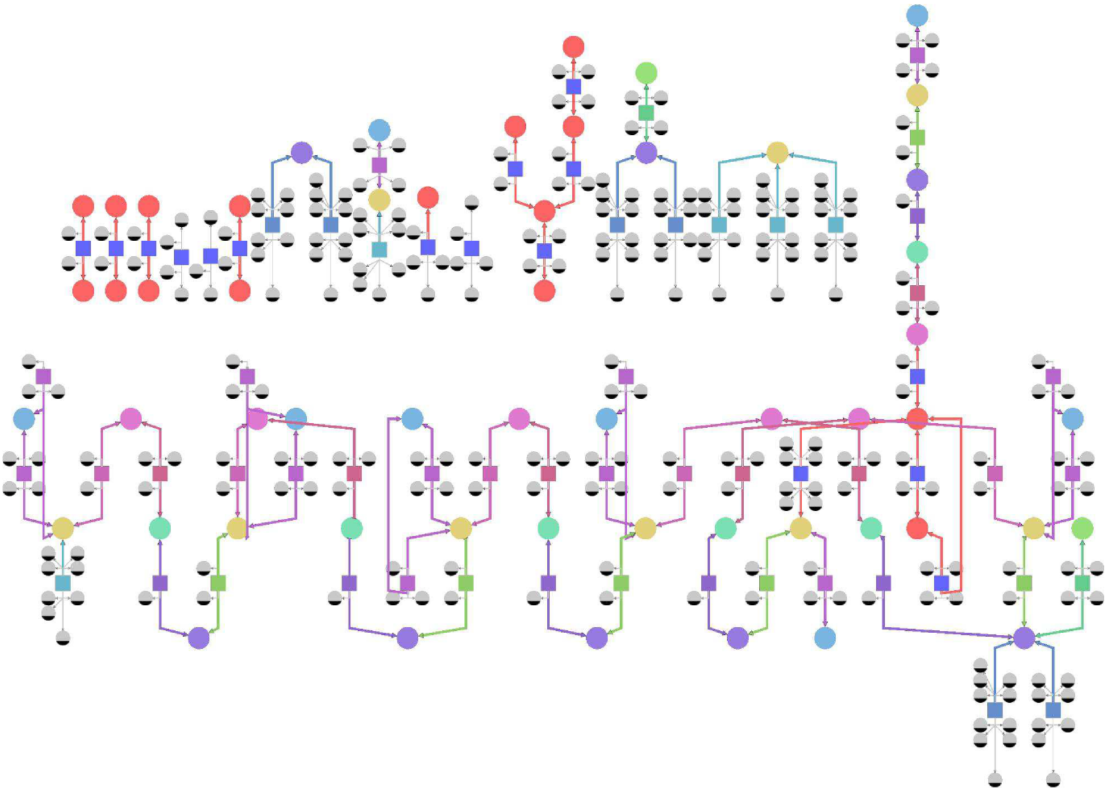


Figure 2. Yarrowia lypolitica fatty acid oxidation model before generalization. Reactions of the specific model are divided into fifteen equivalence classes, represented by different colours. Generally speaking, β-oxidation is a transformation of fatty acyl-CoA (yellow) into dehydroacyl-CoA (violet), then into hydroxyacy fatty acyl-CoA (dark green), 3-ketoacyl-CoA (magenta), and back to fatty acyl-CoA (with a shorter carbon chain); while the specific model describes the same process in more details, specifying those reactions for each of the fatty acyl-CoA species presented in the organisms' cell (e.g. decanoyl-CoA, dodecanoyl-CoA, etc.). This high-level, repetitive structure is obscured by the detail of the individual reactions.

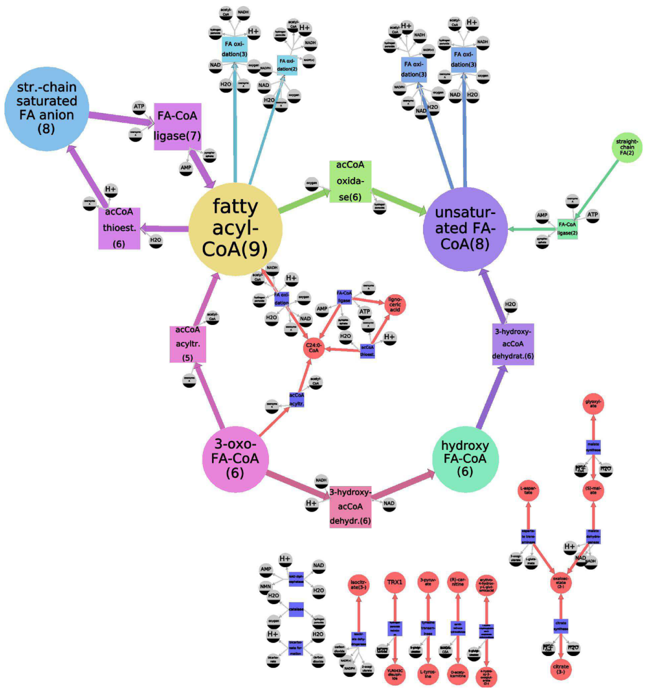


Figure 3. Generalization of the Yarrowia lypolitica fatty acid oxidation model, described as a transformation of fatty acyl-CoA (yellow) into dehydroacyl-CoA (violet), then into hydroxyacy fatty acyl-CoA (dark green), 3-ketoacyl-CoA (magenta), and back to fatty acyl-CoA with a shorter carbon chain. The generalization algorithm identifies equivalent molecular species using an ontology, and groups together reactions that operate on the same abstract species. It finds the greatest generalization the preserves stoichiometry. The generalized model represents quotient species and reactions. For example, the violet dehydroacyl-CoA node is a quotient of hexadec-2-enoyl-CoA, oleoyl-CoA, tetradecenoyl-CoA, trans-dec-2-enoyl-CoA, trans-dodec-2-enoyl-CoA, trans-hexacos-2-enoyl-CoA, trans-octadec-2-enoyl-CoA, and trans-tetradec-2-enoyl-CoA (colored violet in figure 2). In a similar manner, the light-green acyl-CoA oxidase quotient reaction, that converts fatty acyl-CoA (yellow) into dehydroacyl-CoA (violet), generalizes six corresponding light-green reactions of the initial model (figure 2).

MAGRIT Project-Team

6. New Results

6.1. Highlights of the Year

We were invited to present our work on *Impact of Soft Tissue Heterogeneity on Augmented Reality for Liver Surgery* in the TVCG Special Session at SIGGRAPH Asia 2014.

6.2. Matching and 3D tracking

Pose initialization

Automating the camera pose initialization is still a problem in non instrumented environments. Difficulties originate in the possibly large viewpoint changes and lighting variations between the data stored in the model and the current view. One year ago, we began to investigate the use of viewpoint simulation techniques for re-localization within P. Rolin's PhD thesis. We especially consider challenging situations where the current view is distant from the image sequence used for model construction. We here consider scene models built from image sequence using Structure from Motion techniques. A point is then represented by its 3D coordinates and small image patches arising from the images where the point is detected.The underlying idea is to enrich 3D points by descriptors generated from virtual viewpoints chosen away from the learning sequence. For each 3D point of the model, a local image patch is generated from a set of virtual viewpoints, taking into account the local 3D normal and the images of the learning sequence. View synthesis is performed with an affine or an homography model. Though one possible shortcoming of simulation is to generate too many incorrect patches at discontinuities in the scene and thus to degrade the matching step, our preliminary results are very promising [25] and show a noticeable increase of the inlier ratio in the matching stage and an improved stability of the computed pose, especially when homography models are considered. We exhibit many examples where our method successfully computes the camera pose whereas the traditional methods fail.

Current investigations are about the development of scalable solutions for pose computation in large environments with several leverage actions in view. Designing efficient probabilistic techniques for matching and defining strategies based on the geometry of the scene for choosing a reduced set of virtual views are lines of research under investigation for jointly limiting the redundancy and improving the performance of the matching.

Tracking 3D deformable objets

3D augmentation of deformable objects is a challenging problem with many potential applications in computer graphics, augmented reality and medical imaging. Most existing approaches are dedicated to surface augmentation and are based on the inextensibility constraint, for sheet-like materials, or on the use of a model built from representative samples. However, few of them consider in-depth augmentation which is of utmost importance for medical applications. Since the beginning of N. Haouchine's PhD thesis, we have addressed several important limitations that currently hinder the use of augmented reality in the clinical routine of minimally invasive procedures. In collaboration with the SHACRA team, our main contribution is the design and the validation of an augmented reality framework based on a mechanical model of the organ and guided by features extracted and tracked on the video at the surface of the organ [2]. Specific models which best suit the considered organs, such as a vascularized model of the liver, have been introduced in this framework. During this year, we have first performed quantitative evaluation of the method [17]. Promising results were obtained through in-vivo experimentation on a human liver and ex-vivo validation on a porcine liver. In this latter case, artificial tumors were introduced in the liver, thus allowing a quantitative evaluation of the error between the predicted and the actual tumor. These experiments show that localization errors were less than 6mm, and thus below the safety margin required by surgery. To our knowledge, we were the first to produce such evaluation for deformable objects. This work has been extended to augment highly elastic objects in a monocular context [16], whereas previous works were guided by 3D features obtained with a stereo-endoscope. The only parameter involved in the method is the Young's modulus but we show in experiments that a rough estimate of the Young's modulus is sufficient to obtain a good reconstruction. Experiments on computer-generated and real data have shown the effectiveness of the approach. The method is currently restricted to the orthographic projection and its extension to full projective geometry is under investigation.

A bio-mechanical model-based approach has also been considered in the context of tongue tracking in ultrasound images with a view to produce an augmented head for language learning. A crucial issue is the robustness of the tracking due to the strong speckle noise in ultrasound (US) data. Here, a small number of points are used to guide the model. Selection of feature points is based on the uncertainty associated to the tracked points and on spatial constraints. This model has proven to be especially efficient in the case of non uniform and fast movements [19].

Use of AR in educational sciences
In collaboration with the Ecole supérieure du professorat et de l'éducation and the PErSEUs laboratory at Université de Lorraine, we designed an inquiry-based AR learning environment (AIBLE) for teaching and learning astronomy in primary school (children of 8-11 years old). The novelty of this environment is the combination of Inquiry Based Sciences Education principles and didactics principles (here of astronomy) with AR capabilities. In this context, a GPL-licensed software called AIBLE-AstroAR has been developed based on the ARToolkit library. This software basically consists of a tangible user interface, which allows the children to move virtual celestial objects "as for real" and investigate in order to find origins of Moon phases evolution, alternation of day and night, seasons and Moon/Sun eclipses.

Last year, a study has been carried out to compare AIBLE with a physical model traditionally used in primary school. This study indicated that AIBLE significantly enhances learning compared to classical support. During this year, we performed further investigation with a larger panel of children to assess which characteristics of the environment facilitate learning [14]. Analyses of the marker positions as moved by the children indicated that AIBLE really facilitates heuristic investigation, which fosters consciousness of the origin of astronomical phenomena. This work provides new opportunities for teachers to identify solving problem strategies initiated by learners. These results also contribute to the understanding of the ways through which AR can be used in formal teaching curricula in K-12 schools.

6.3. Image-based modeling

Modeling vasculature for real time simulation
One of our objectives to benefit interventional neuroradiology is to offer a patient-based interactive simulator to the interventional radiologists. Our contributions address vasculature modeling from patient data, namely 3D rotational angiography (3DRA) volumes. During Ahmed YUREIDINI's PhD thesis (2010-2014), a new model was developed consisting of a tree of local implicit blobby models.

We've been collaborating with SHACRA Inria project-team (Lille-Nord Europe) and the Department of Interventional Neuroradiology from Nancy University Hospital, in the context of the SOFA-InterMedS Inria Large-Scale Initiative. Ahmed YUREIDINI defended his PhD thesis in May this year with highest honors [9]. In particular, a detailed study was made to compare our tree of local implicits with triangular meshes in a view to model synthetic shapes as well as vasculatures from patient data. Increased performances with regard to processing speed, numerical stability and realism of the behavior were demonstrated.

Tools reconstruction for interventional neuro-radiology
Minimally invasive techniques impact surgery in such ways that, in particular, an imaging modality is required to maintain a visual feedback. Live X-ray imaging, called fluoroscopy, is used in interventional neuroradiology. Such images are very noisy, and cannot show but the vasculature and no other brain tissue. In particular, since at most only two projective fluoroscopic views are available, containing absolutely no depth hint, the 3D shape of the micro-tool (guidwire, micro-catheter or micro-coil) can be very difficult, if not impossible to infer, which may have an impact on the clinical outcome of the procedure.

In collaboration with GE Healthcare, we aim at devising ways to reconstruct the micro-tools in 3D from fluoroscopy images. Charlotte Delmas has been working as a PhD Cifre student on this subject since April 2013. A solution in a two-view reconstruction context was proposed this year based on the extraction of the guide-wire as a skeleton in the images. The large stereo basis (views are almost orthogonal) and the segmentation errors (such as both missing parts and spurious segments in the skeleton) make the reconstruction especially difficult. The skeletons are subdivided in simple curves that are matched to build all corresponding potential 3D curves. These curves are nodes in a graph whose edge weights express a connection cost that takes into account both distance and orientation at the curves extremities. The solution 3D curve is provided by following the path of minimal cost in the graph. This algorithm demonstrated very good reconstruction results on synthetic and phantom data. A paper on this subject has been accepted for publication at SPIE Medical Imaging 2015.

Patient-specific heart valve modeling

Many pathologies damage heart valve anatomy producing undesired backflow, or regurgitation, decreasing cardiac efficiency and potentially leading to heart failure if left untreated. Such cases could be treated by surgical repair for the valve. However it is technically difficult and outcomes are highly dependent upon the experience of the surgeon: he must essentially predict the displacement and deformation of complex valve leaflets and supporting structures. One way to facilitate the repair is to simulate the mechanical behavior of the pathological valve with patient-specific data. This is the objective of Pierre-Frédéric Villard's one-year CNRS delegation in the Harvard Bio-robotics Laboratory (HBL). During the initial three first months of the sabbatical leave, various tasks have been performed: i) Study of the physiology of pathological valve behavior with medical experts. Following anatomical book reading and medical expert interviews the anatomy and the physiology are now understood. ii) Evaluation of HBL material for 4D ultrasound segmentation. HBL has previously developed a method to extract mitral valve geometry from a home-made high temporal resolution 3D ultrasound and iii) Automatic segmentation of a Mitral Valve microCT to feed a biomechanical model. A method to semi-automatically segment the leaflet-chordae set has been developed.

6.4. Parameter estimation

Metrologic performance assessment in experimental mechanics

A problem of interest in experimental solid mechanics is strain map estimation on the surface of a specimen subjected to a load or a tensile test. One of the available approaches is based on images of a pseudo-periodic grid transfered on the surface of the specimen. Sensor noise is a major source of uncertainty in the strain map, and quantifying the propagation of the sensor noise to the measured strain components is a major problem when metrological performances are in view. We have proposed in [12] a study of the mathematical properties of the popular method based on windowed Fourier analysis, under a Gaussian white noise assumption. In the case of a more realistic signal-dependent, heteroscedastic noise, we have quantified in [10] (see also [15], [26]) the trade-off between the noise amplitude, the measurement resolution and the spatial resolution of the method. We have also investigated image stacking for noise reduction. While averaging a serie of images is certainly the most basic option to reduce the noise, it is not effective for studying grid images under a high magnification factor, because of unavoidable residual vibrations carried for instance by concrete floor slabs. We have shown in [13] that, while these vibrations indeed blur grid images, they still permit to reduce the noise amplitude in the displacement and strain maps.

Sensor noise measurement.

While searching for a low-cost on-the-fly estimation of the sensor parameters based on a serie of grid images (thus with no need of changing the experimental setting), we have proposed in [11] an algorithm which is able to deal with the vibrations biasing the estimations. More generally, we have investigated in [21] the problem of sensor parameter estimation from a series of images, under light flickering and vibrations. Light flickering is indeed a natural assumption for indoor artificial lights. It is also involved by slight variations in the opening time of a mechanical shutter. We have proposed a model of the pixel intensity based on a Cox process, together with an algorithm which, taking benefit of flickering, gives an estimation of every sensor parameter, namely the gain, the readout noise, and the offset.

Image driven simulation

In the IDeaS ANR project we propose to target Image-driven simulation, applied to interventional neuroradiology: a coupled system of interactive computer-based simulation (interventional devices in blood vessels) and on-line medical image acquisitions (X-ray fluoroscopy). The main idea is to use the live X-ray images as references to continuously refine the parameters used to simulate the blood vessels and the interventional devices (micro-guide, micro-catheter, coil).

Our guideline is to follow a sequential statistical filtering approach to fuse such heterogeneous data. This approach first calls for an improved knowledge of the statistical behavior of the simulation, which we addressed in the past year through experimental studies. We described our experimental setup in [20], which, in particular uses high speed stereo reconstruction to be able to study non quasi-static effects. Preliminary measures of the catheter speed during stick and slip transitions back up our conviction that quasi-static mechanical models fail to simulate such rapid motions of the tool. Our on-going analysis of the simulation sensitivity to mechanical parameters also sets forward friction as critical for high-fidelity simulation.

MAIA Project-Team

6. New Results

6.1. Highlights of the Year

- Two Research Fellow have been recruited with a focus on Service Robotics: Serena Ivaldi (CR2) and Francis Colas (CR1).
- The paper entitled : Exploiting Separability in Multiagent Planning with Continuous-State MDPs Jilles Dibangoye, Christopher Amato, Olivier Buffet, François Charpillet won the best paper award at AAMAS'2014, the international conference on autonomous agents and multi-agents.
- Jilles Dibangoye got an Assistant Professor position at INSA Lyon.

BEST PAPER AWARD :

[12] **13th International Conference on Autonomous Agents and Multiagent Systems**. J. S. DIBANGOYE, C. AMATO, O. BUFFET, F. CHARPILLET.

6.2. Decision Making

6.2.1. Complexity Analysis of Exact Dynamic Programming Algorithms for MDPs

Participant: Bruno Scherrer.

Eugene Feinberg and Jefferson Huang are external collaborators from Stony Brooks University.

Following last year's work on the strong polynomiality of Policy Iteration, we show that the number of arithmetic operations required by any member of a broad class of optimistic policy iteration algorithms to solve a deterministic discounted dynamic programming problem with three states and four actions may grow arbitrarily. Therefore any such algorithm is not strongly polynomial. In particular, the modified policy iteration and λ-policy iteration algorithms are not strongly polynomial. This work was published in the *Operations Research Letters* [4].

6.2.2. Analysis of Approximate Dynamic Programming Algorithms for MDPs

Participants: Bruno Scherrer, Manel Tagorti.

Matthieu Geist is an external collaborator from Supélec.

In [40], we consider LSTD(λ), the least-squares temporal-difference algorithm with eligibility traces algorithm proposed by Boyan (2002). It computes a linear approximation of the value function of a fixed policy in a large Markov Decision Process. Under a β-mixing assumption, we derive, for any value of $\lambda \in (0, 1)$, a high-probability estimate of the rate of convergence of this algorithm to its limit. We deduce a high-probability bound on the error of this algorithm, that extends (and slightly improves) that derived by Lazaric et al. (2012) in the specific case where $\lambda = 0$. In particular, our analysis sheds some light on the choice of λ with respect to the quality of the chosen linear space and the number of samples, that complies with simulations. This work was presented at the National JFPDA conference [34].

In the context of infinite-horizon discounted optimal control problem formalized by Markov Decision Processes, we focus on several approximate variations of the Policy Iteration algorithm: Approximate Policy Iteration (API), Conservative Policy Iteration (CPI), a natural adaptation of the Policy Search by Dynamic Programming algorithm to the infinite-horizon case (PSDP), and the recently proposed Non-Stationary Policy Iteration (NSPI). For all algorithms, we describe performance bounds with respect the per-iteration error ϵ, and make a comparison by paying a particular attention to the concentrability constants involved, the number of iterations and the memory required. Our analysis highlights the following points: 1) The performance guarantee of CPI can be arbitrarily better than that of API, but this comes at the cost of a relative—exponential in $\frac{1}{\epsilon}$—increase of the number of iterations. 2) PSDP_∞ enjoys the best of both worlds: its performance guarantee is similar to that of CPI, but within a number of iterations similar to that of API. 3) Contrary to API that requires a constant memory, the memory needed by CPI and PSDP is proportional to their number of iterations, which may be problematic when the discount factor γ is close to 1 or the approximation error ϵ is close to 0; we show that the NSPI algorithm allows to make an overall trade-off between memory and performance. Simulations with these schemes confirm our analysis. This work was presented at this year's international conference on Machine Learning (ICML) [28].

Finally, we consider Local Policy Search, that is a popular reinforcement learning approach for handling large state spaces. Formally, it searches locally in a parameterized policy space in order to maximize the associated value function averaged over some predefined distribution. The best one can hope in general from such an approach is to get a local optimum of this criterion. The first contribution of this article is the following surprising result: if the policy space is convex, *any* (approximate) *local optimum* enjoys a *global performance guarantee*. Unfortunately, the *convexity* assumption is strong: it is not satisfied by commonly used parameterizations and designing a parameterization that induces this property seems hard. A natural solution to alleviate this issue consists in deriving an algorithm that solves the local policy search problem using a boosting approach (constrained to the convex hull of the policy space). The resulting algorithm turns out to be a slight generalization of conservative policy iteration; thus, our second contribution is to highlight an original connection between local policy search and approximate dynamic programming. This work was presented at this year's European conference on Machine Learning (ECML) [27].

6.2.3. Adaptive Management with POMDPs

Participants: Olivier Buffet, Jilles Dibangoye.

Samuel Nicol and Iadine Chadès (CSIRO) are external collaborators.

In the field of conservation biology, adaptive management is about managing a system, e.g., performing actions so as to protect some endangered species, while learning how it behaves. This is a typical reinforcement learning task that could for example be addressed through Bayesian Reinforcement Learning.

During Samuel Nicol's visit, the main problem we have studied is how to manage company inspections to deter these companies from adopting dangerous behaviors. This was modeled as a particular Stackelberg game, where N companies benefit from acting badly as long as they are not caught by inspections, and where 1 government agency has to decide which companies to inspect given a limited budget. The expected result is a stochastic strategy (randomly deciding which companies to inspect, with probabilities that depend on the benefits/losses of both types of players). We are working on exploiting particular features of this computationally complex problem to make it more tractable.

6.2.4. Solving decentralized stochastic control problems as continuous-state MDPs

Participants: Jilles Dibangoye, Olivier Buffet, François Charpillet.

External collaborators: Christopher Amato (MIT).

Decentralized partially observable Markov decision processes (DEC-POMDPs) are rich models for cooperative decision-making under uncertainty, but are often intractable to solve optimally (NEXP-complete), even using efficient heuristic search algorithms.

State-of-the-art approaches relied on turning a Dec-POMDP into an equivalent deterministic MDP —whose actions at time t correspond to a vector containing one decision rules (/instantaneous policy) per agent— typically solved using a heuristic search algorithm inspired by A*. In recent work (IJCAI'13), we have identified a sufficient statistic of this MDP —an *occupancy state*, i.e., a probability distribution over possible states and joint histories of the agents— and demonstrated that the value function was piecewise-linear and convex with respect to this statistic. This brings us in the same situation as POMDPs, allowing to generalize the value function from one occupancy state to another and to propose much faster algorithms (also using efficient compression methods).

This year, we have further progressed on this line of research.

- A journal paper has been submitted that presents the "occupancy MDP" approach in details.
- In the case of Network-Distributed POMDPs, a particular setting where the relations between agents follow a fixed network topology, we have shown that the value function could be decomposed additively with one value function per neighborhood. This work has been presented at AAMAS'2014 [12], receiving the conference's best paper award.
- To further scale up the resolution of Dec-POMDPs, we have proposed multiple approximations techniques that can be combined and allow controlling error bounds. This work has been presented at ECML'2014 [13].

6.2.5. Learning Bad Actions

Participant: Olivier Buffet.

Jörg Hoffmann, former member of MAIA, Michal Krajňanský (Saarland University), and Alan Fern (Oregon State University) are external collaborators.

In classical planning, a key problem is to exploit heuristic knowledge to efficiently guide the search for a sequence of actions leading to a goal state.

In some settings, one may have the opportunity to solve multiple small instances of a problem before solving larger instances, e.g., trying to handle a logistics problem with small numbers of trucks, depots and items before moving to (much) larger numbers. Then, the small instances may allow to extract knowledge that could be reused when facing larger instances. Previous work shows that it is difficult to directly learn rules specifying which action to pick in a given situation. Instead, we look for rules telling which actions should not be considered, so as to reduce the search space. But this approach requires considering multiple questions: What are examples of bad (or non-bad) actions? How to obtain them? Which learning algorithm to use?

A first algorithm (with variants) has been proposed that learns rules for detecting (supposedly) bad actions. It has been empirically evaluated, providing encouraging results, but also showing that different variants will perform best in different settings. This algorithm has been presented at ECAI'2014 [24], and has participated in the learning track of the international planning competition in 2014 (http://ipc.icaps-conference.org/).

6.3. Ambiant Intelligence And Robotic Systems

6.3.1. Adaptation of autonomous vehicle traffic to perturbations

Participants: Mohamed Tlig, Olivier Buffet.

Olivier Simonin, a former member of the MAIA team, is an external collaborator from INSA-Lyon.

The aim of the European project InTraDE is to propose more efficient ways to handle containers in seaports through the use of IAVs (Intelligent Autonomous Vehicles).

In his PhD thesis, Mohamed Tlig considers the displacements of numerous such IAVs whose routes are a priori planned by a supervisor. However, in such a large and complex system, different unexpected events can arise and degrade the traffic: failure of a vehicle, human mistake while driving, obstacle on roads, local re-planning, and so on.

In 2013, we have started looking at improving vehicle flows in complete road networks. In particular, we have proposed an approach that allows multiple flows of vehicles to cross an intersection without stopping, allowing to reduce delays as well as energy consumption. This has led to a publication in ICALT-14 [30], with more details in a research report [41].

This year, we have made a further step by coordinating the controller agents located in each of the network's intersections. More precisely, they are constrained to let the vehicles alternate at the same frequency —at the expense of potentially reducing the maximum flow of some roads— and a distributed algorithm offsets these "signals" so as to optimize either the energy consumption, or the time spent in the network. This tends to induce "green waves" wherever possible, i.e., to prevent vehicles from having to slow down before a traffic light. This work has been presented at ECAI-14 [31].

6.3.2. Platooning: safe and precise virtual hooking mechanism or automated vehicles

Participants: Jano Yazbeck, Alexis Scheuer, François Charpillet.

Among the several goals that we were trying to achieve in InTraDE, we were interested in platooning too. In her PhD thesis, Jano Yazbeck considers Platooning as a technique that aims at steering , safely and precisely, a train of vehicles along a path generated by a leader which can be driven by a human. Thus the trajectory is unknown to the followers. Platooning is considered in this project in order to move containers efficiently from the discharge zones of ships to the storage areas.

To obtain a safe and precise platooning, we aim at controlling the longitudinal and lateral behaviors of each vehicle of the platooning. On the one hand, the longitudinal controller computes a longitudinal velocity (or acceleration) which avoids collisions between vehicles by maintaining a safe inter-distance between each couple of successive vehicles. On the other hand, the lateral controller computes an angular velocity or a steering angle so that the vehicle follows precisely the leader's path. These two controllers can be decoupled and computed separately when the convoy moves at a low velocity.

This year, we proposed a platooning algorithm based on a near-to-near decentralized approach which has been published at ICRA 2014 [32]. In this approach, each vehicle estimates and memorizes on-line the path of its predecessor as a set of points. After choosing a suitable position to aim for, the follower estimates on-line the predecessor's path curvature around the selected target. Then, based on a heuristic search, it computes an angular velocity using the estimated curvature. The optimization criteria used in this work allows the robot to follow its predecessor's path without oscillation while reducing the lateral and angular errors.

In october, Jano Yazbeck defended her Phd Thesis [2].

6.3.3. Map Matching

Participant: François Charpillet.

This work [8] has been realized during the Intrade Projet with Maan Badaoui from Lille University. It addresses an important issue for intelligent transportation system, namely the ability of vehicles to safely and reliably localize themselves within an a priori known road map network. For this purpose, we have proposed an approach based on hybrid dynamic bayesian networks enabling to implement in a unified framework two of the most successful families of probabilistic model commonly used for localization: linear Kalman filters and Hidden Markov Models. The combination of these two models enables to manage and manipulate multi-hypotheses and multi-modality of observations characterizing Map Matching problems and it improves integrity approach. Another contribution of the paper is a chained-form state space representation of vehicle evolution which permits to deal with non-linearity of the used odometry model. Experimental results, using data from encoders' sensors, a DGPS receiver and an accurate digital roadmap, illustrate the performance of this approach, especially in ambiguous situations.

6.3.4. Multi-Camera Tracking in Partially Observable Environment

Participants: Arsène Fansi Tchango, Olivier Buffet, Vincent Thomas, Alain Dutech.

Fabien Flacher (Thales ThereSIS) is an external collaborator.

In collaboration with Thales ThereSIS - SE&SIM Team (Synthetic Environment & Simulation), we focus on the problem of following the trajectories of several persons with the help of several controllable cameras. This problem is difficult since the set of cameras cannot simultaneously cover the whole environment, since some persons can be hidden by obstacles or by other persons, and since the behavior of each person is governed by internal variables which can only be inferred (such as his motivation or his hunger).

The approach we are working on is based on (1) the HMM (Hidden Markov Models) formalism to represent the state of the system (the persons and their internal states), (2) a simulator provided and developed by Thales ThereSIS, and (3) particle filtering approaches based on this simulator. Since activity and location depend on each other, we adopt a Simultaneous Tracking and Activity Recognition approach (STAR) as presented in current state-of-the-art approaches.

A first novelty lies in the use of a complex behavioral simulator. In a single-target setting, we demonstrated that it allows inferring the behavior of a complex individual, even in case of long periods of occlusions (when cameras do not cover the trajectory of the target). This idea led to publications in AAMAS-14 [16], STAIRS-14 [18], and ECAI-14 [17].

A remaining issue is to find tractable algorithms for efficiently tracking multiple targets simultaneously, which requires using a factored particle filter (with one distribution per target). To that end, we use a Joint Probabilistic Data Association Filter with two key ingredients. The first ingredient is a particular model of dynamics that largely decouples the evolution of several targets, and turns out to be very natural to apply (which has led already to a publication in Fusion-14 [19]). Then, the factorization *a priori* implies, for a given target, simulating each of its particles with each particle of each other target (which leads to a huge number of simulations). The second proposed ingredient is to simulate each particle of a given target only with a small number of “representatives” of each other target (and then, because more particles are produced than needed, a selection/resampling step is required).

6.3.5. Emergence et Developmental Learning

Participants: Alain Dutech, Matthieu Zimmer.

Yann Boniface (CORTEX, Loria) is an external collaborator

Following our ongoing work on using reinforcement learning for the control of redundant continous robotic systems, we explore how learning such complex tasks can benefit from a developmental approach, following some line of work already tested in robotics [50].

“Emergence”, on of the key concepts grounding this work, has been presented – from an artificial intelligence perspective – and discussed with researchers from other fields. This lead to fruitful exchanges and a chapter in a bookdedicated to the dual aspects of (human gestures) : appearance and emergence [36]. “Developmental Learning” was also the main subject of a seminar in Lyon in which Alain Dutech has been invited [47].

More concretely, we have developed several algorithms which mix artificial neural networks (like Dynamic Self-Organizing Maps or Reservoir Computing Network) with reinforcement learning mechanisms in order to build simple artificial systems that are *autnomous* and that learn without any *exogeneous* intervention from an external being. This work, initiated through two master thesis, is now the central topic ot the PhD of Matthieu Zimmer, started in october 2014.

6.3.6. Online Evolutionary Learning

Participants: Amine Boumaza, François Charpillet, Iñaki Fernandèz.

Evolutionary Robotics (ER) deals with the design of agent behaviors using artificial evolution. Within this framework, the problem of learning optimal decision functions (or controllers) is treated as a policy search problem in the parameterized space of candidate policies. In this work we are interested in learning optimal behaviors for swarm of mobile agents online (while solving the task). We adopt an online onboard distributed view [56], [48] and consider the learning process as executed at the agents’ level in a decentralized way. This kind of algorithms raises several questions concerning the usefulness of selection pressure (partial views of population, noisy fitness values, etc.).

We studied the impact of task-driven selection pressures in on-line distributed ER for swarm behavior learning. We proposed a variant of the mEDEA [45] algorithm in which we added a selection operator, in a task-driven scenario. We evaluated four selection methods that induce different intensity of selection pressure in a multi-robot navigation with obstacle avoidance task and a collective foraging task.

Experiments showed that a small intensity of selection pressure is sufficient to rapidly obtain good performances on the tasks at hand. We introduced different measures to compare the selection methods, and show that the higher the selection pressure, the better the performances obtained, especially for the more challenging food foraging task. This research was presented at the 13th International Conference on the Synthesis and Simulation of Living Systems [21].

6.3.7. Frailty evaluation and Fall detection

Participants: Amandine Dubois, François Charpillet, Thomas Moinel, Maxime Rio.

This work is related to the IPL PAL and Satelor project and is related to Personal Assistant Living (PAL) for elderly people with loss of autonomy.

- Clinical evaluation of frailty in the elderly is the first step to decide the degree of assistance that elderly people require. No standard tests exist to detect the level of frailty, each clinician chooses his protocol among existing tools. There are clinical tests as *Tinetti test*, *Timed Up and Go* test for evaluating the degree of dependance and the frailty of elderly people. These tests consist in asking a person to realize exercises simulating movements of daily life. The physician evaluates the quality of gait and the balance of the patient. These tests are often used but, the disadvantage is that the final verdict relies primarily on a subjective opinion. The aim of our work is to provide new objective criteria to refine the elderly frailty quantification. We base ourselves on the frailty definition of Fried *et al* as being a clinical syndrome in which three or more of the following criteria are present: unintentional weight loss, self reported exhaustion, weakness (with regards to grip strength), slow walking speed and low physical activity. From this definition, we have defined two axis of development to evaluate the frailty of a person: Sensor based Activity recognition with the aim to follow and report daily life activities in order to detect evolution that coud reveal increased frailty [1], gait analysis in order to assess gait pattern and their evolution over time [14].
- An other PAL research domain, which is related to activity recognition, has attracted our attention: fall detection. Falls in the elderly is a major public health problem because of their frequency and their medical and social consequences. One of our objectives is to design an automatic system to detect fall at home, which in its final version will be made up of a network of RGB-D sensors, some of them being mobile embedded a wheel mobile robot.

The main contribution of this work has been to design a simple but robust method based on the identification and tracking of the center of mass of people evolving in an indoor environment through a RGB-D camera. Using a simple Hidden Markov Model whose observations are the position of the center of mass, its velocity and the general shape of the body, we have shown that we can surprisingly monitor the activity of a person with high accuracy, detect falls with very good accuracy without false positives and also measure some interesting parameter such as speed of gait, length of steps, etc. An experimental study, that is reported in [46], has been driven in our smart apartment lab. 26 subjects were asked to perform a predefined scenario in which they realized a set of eight postures. 2 hours of video (216 000 frames) were recorded for the evaluation, half of it being used for the training of the model. The system detected the falls without false positives. This result encourages us to use this system in real situation for a better study of its efficiency. Therefore, we started this year an experimention in a room of a follow-up care and rehabilitation facility (OHS) in Nancy. "Office d'Hygiène Sociale" (OHS) is an association under the law of 1901. It supports nearly 800 people over 60 years and nearly 1,000 children and adults with disabilities. The association manages 26 facilities (40% health field, 40% medical-social field and 20% social field) and employs more than 1,500 professionals.

6.3.8. Posture recognition with a Depth camera

Participants: Abdallah Dib, François Charpillet, Xuan Nguyen, Alain Filbois [SED].

In this research line, we focus our contribution on improving model-based approaches that use a population-based stochastic framework for full human body tracking using monocular depth camera. One of the major challenges in human tracking is the high-dimensional state spaces. To address this problem, we propose a tracking algorithm based on APF and CMA-ES. While APF has been widely applied for human tracking in RGB and depth images, the application of CMA-ES to human tracking is still limited. Yet, CMA-ES shares many similar ideas with APF and can be exploited to improve the performance of APF. Our key idea is to update the covariance matrix for sampling particles at each layer of APF, using a subset of best particles, an idea inspired from CMA-ES. The resulting algorithm is shown to greatly reduce the number of particles required for successful tracking. In the absence of image features such as texture or color, existing likelihood models for human tracking in depth images are often built by computing distances between data points and model points sampled on the surface of the human body model. When human body parts are close or when severe self-occlusions are present, these models fail to capture good pose hypotheses. As a result, existing approaches are unable to track a broad range of human motions. To deal with this issue, we propose a likelihood model which is based on comparing observed depth images and rendered depth images obtained by classic rendering techniques. Combining with our tracking algorithm, the proposed likelihood model has been shown to be effective when tracking under severe self-occlusions. To the best of our knowledge, our approach is the first model-based one that uses a population-based stochastic framework able to track full human body with non-frontal and unusual poses, using monocular depth camera.

6.3.9. Pressure sensing floor

Participants: Mihai Andries, François Charpillet, Olivier Simonin.

The use of floor-sensors in ambient intelligence contexts began in the late 1990's, with projects like ORL active floor, the Magic carpet by Paradiso *et al.*, and the smart floor by Orr *et al.* These floors were, later on, integrated in smart environments, aimed at delivering assistance services like continuous diagnosis of users' health. According to the literature there are currently at least 6 main types of floor pressure sensing technologies: binary switches, piezoelectric, load cells, capacitive, polymer thick film (PTF), and photo interrupter sensors. Most of presented solutions extract a set of features for their tracking and identification task. Recently, sensing floors products like the SensFloor (a floor network of capacitive proximity sensors), Capfloor (a network of capacitive sensors), Elsi® smart floor (http://www.elsitechnologies.com) and FloorInMotion (Tarkett France) started being commercialized by companies, mainly for the senior care industry.

We have ourselves developed a sensing floor. This load-sensing floor is composed of square tiles, each equipped with two ARM processors (Cortex m3 and a8), 4 load cells, and a wired connection to the four neighboring cells. Each tile has 16 light-emitting diodes which provide visual feedback. The processing units were manufactured by Hikob [0]. This prototype was originally designed as a medium of interaction for robots with distributed control, in an ant-like fashion. The computing unit available on each tile can register a virtual pheromone trace, that can then be transmitted to other robots, using either wired or wireless communication. In a different perspective, the sensing-floor acts merely as a sensor for an ambient intelligence. Using the magnetometer embedded on the processing unit of the tile, each tile can detect disturbances in its surrounding magnetic field, that can be caused by the presence of robots. Each tile also has an embedded accelerometer, that allows it to detect shocks that can be caused by objects or humans falling on the ground.

Several functionalities have been implemented this year on this prototype floor, including weight measurement, fall detection, footstep tracking and activity recognition. We also implemented heuristic real-time multi-user localisation (without user identification) in an indoor setting using this prototype floor.

6.3.10. Living assistant Robot

Participants: François Charpillet, Nicolas Beaufort, Abdallah Dib.

[0] http://www.hikob.com/

With LAR (**living AssistanT Robot**), a PIA projet which started in March, Abdallah Dib joined our team for a PhD. His work is about the development of a low cost navigation system for a robot evolving in an indoor environment. The main issue of his work is to design a Simultaneous Localisation and Mapping algorithm working in a dynamic environment in which people are moving. This is very challenging if we restrict the sensing capabilities of the robot with low cost sensors such as RGB-D camera. An important service we expect the robot to achieve, is realizing similar services as the one we described below: fall detection, activity recognition. This year first result have been published [11]. A feature based visual SLAM method that uses chamfer distance to estimate the camera motion from RGB-D images has been presented. The method does not require any matching which is an expensive operation and always generates false matching that affects the estimated camera motion. Our approach registers the input image iteratively by minimizing the distance between the feature points and the occupancy grid using a distance map. We demonstrate with real experiments the capability of the method to build accurate 3D map of the environment with a hand-held camera. While the system was mainly developed to work with RGB-D camera, occupancy grid representation gives the method the ability to work with various types of sensors, we show the capacity of the system to construct accurate 2D maps using telemeter data. We also discuss the similarities between the proposed approach and the traditional ICP algorithm.

6.3.11. Exploring an unknown environment with a team of mobile robots

Participants: François Charpillet, Olivier Simonin, Nassim Kaldé.

This work is the continuation of the work realized during the ANR Cart-O-matic (2010 to 2013). We address, here, the problem of efficient allocation of the navigational goals in the multi-robot exploration of unknown environment. Goal candidate locations are repeatedly determined during the exploration. Then, the assignment of the candidates to the robots is solved as the task-allocation problem. A more frequent decision-making may improve performance of the exploration, but in a practical deployment of the exploration strategies, the frequency depends on the computational complexity of the task-allocation algorithm and available computational resources. Therefore, this year, we have proposed an evaluation framework to study exploration strategies independently on the available computational resources. A comparison of the selected task-allocation algorithms deployed in multi-robot exploration has been done and published with Jan Faigl from Czech Technical University in Prague in the framework of the PHC project MACOREX.

An other point that is addressed by Nassim Kaldé is to consider the same problem but with dynamical environment in particular populated with human beings. First results of Nassim Kalde have been published in JFSMA'14 [33]. He published too the work done during his Master thesis [23].

6.4. Understanding and mastering complex systems

6.4.1. Adaptive control of a complex system based on its multi-agent model

Participant: Vincent Chevrier.

Laurent Ciarletta (Madynes team, LORIA) is an external collaborator.

Complex systems are present everywhere in our environment: internet, electricity distribution networks, transport networks. These systems have as characteristics: a large number of autonomous entities, dynamic structures, different time and space scales and emergent phenomena. This thesis work is centered on the problem of control of such systems. The problem is defined as the need to determine, based on a partial perception of the system state, which actions to execute in order to avoid or favor certain global states of the system. This problem comprises several difficult questions: how to evaluate the impact at the global level of actions applied at a global level, how to model the dynamics of an heterogeneous system (different behaviors issue of different levels of interactions), how to evaluate the quality of the estimations issue of the modeling of the system dynamics.

We propose a control architecture based on an "equation-free" approach. We use a multi-agent model to evaluate the global impact of local control actions before applying the most pertinent set of actions.

Associated to our architecture, an experimental platform has been developed to confront the basic ideas or the architecture within the context of simulated "free-riding" phenomenon in peer to peer file exchange networks. We have demonstrated that our approach allows to drive the system to a state where most peers share files, despite given initial conditions that are supposed to drive the system to a state where no peer shares. We have also executed experiments with different configurations of the architecture to identify the different means to improve the performance of the architecture.

This work helped us to better identify [26] the key questions that rise when using the multi-agent paradigm in the context of control of complex systems, concerning the relationship between the model entities and the target system entities.

6.4.2. Multi Modeling and multi-simulation

Participants: Vincent Chevrier, Christine Bourjot, Benjamin Camus, Julien Vaubourg.

Laurent Ciarletta and Yannick Presse (Madynes team, LORIA) are external collaborators.

Laurent Ciarletta is the co-advisor of the thesis of Julien Vaubourg.

Complex systems generally require to use different points of view (abstraction levels) at the same time on the system in order to capture and to understand all the dynamics and the complexity. Being made of different interacting parts, a model of a complex system also requires simultaneously modeling and simulation (M&S) tools from different scientific fields.

We proposed the AA4MM meta-model [54] that solves the core challenges of multimodelling and simulation coupling in an homogeneous perspective. In AA4MM, we chose a multi-agent point of view: a multi-model is a society of models; each model corresponds to an agent and coupling relationships correspond to interaction between agents.

This year we progress in the definition of multi-level modeling [42]. We identified several facets of multi-level modeling and implemented then as different kinds of interactions in AA4MM framework. We progressed on the specification of the meta-model which helped to define a modeling environment.

In the MS4SG projet which involves MAIA, Madynes and EDF R&D on smart-grid simulation, we developed a proof of concepts for a smart-appartment case [10].

6.4.3. Cellular automata as a foundation of complex systems

Participant: Nazim Fatès.

Our research on emergent collective behavior focuses on the analysis of the robustness of discrete models of complex systems. We ask to which extent systems may resist to various perturbations in their definitions. We progressed in the knowledge of how to tackle this issue in the case of cellular automata (CA) and multi-agent systems (MAS).

We proposed an extended version of our survey on asynchronous cellular automata [3].

In collaboration with colleagues from India, we proposed a complete characterisation of the reversibility of the set of the 256 Elementary Cellular Automata with asynchronous updating [29]. These rules are known to be diffcult to study in all generality and it is interesting to notice that here, asynchronism is an aid rather than an obstacle to analyse the behaviour of the systems.

With Henryk Fukś (Brock Univ., Canada), we proposed a mathematical analysis of the second-order phase transitions that are observed in the most simple asynchronous cellular automata [22].

Our work on the classification of cellular automata was presented in the AUTOMATA'14 conference and is now the topic of a collaboration with L. Gerin (École Polytechnique) [44], [20].

We are currently participating to the edition of the first book devoted to probabilistic cellular automata and to a special issue of the French-speaking journal *Technique et Science Informatique* (Lavoisier editors).

6.4.4. Revisiting wavefront construction with collective agents: an approach to foraging.

Participants: François Charpillet, Olivier Simonin.

We consider here [7], the problem of coordinating a team of agents that have to collect disseminated resources in an unknown environment. We are interested in approaches in which agents collectively explore the environment and build paths between home and resources. The originality of our approach is to simultaneously build an artificial potential field (APF) around the agents' home while foraging. We propose a multi-agent model defining a distributed and asynchronous version of Barraquand et al. Wavefront algorithm. Agents need only to mark and read integers locally on a grid, that is, their environment. We prove that the construction converges to the optimal APF. This allows the definition of a complete parameter-free foraging algorithm, called c-marking agents. The algorithm is evaluated by simulation, while varying the foraging settings. Then we compare our approach to a pheromone-based algorithm. Finally, we discuss requirements for implementation in robotics.

MAMBA Team

6. New Results

6.1. Highlights of the Year

Benoît Perthame was invited as plenary speaker for the International Congress of Mathematicians ICM 2014 (Seoul, http://www.icm2014.org), that attracted more than 5000 participants. This is the first time that a mathematician working in mathematics applied to biology was invited at ICM, which is the most prestigious conference for mathematicians of all fields. This represents a consecration both for Benoît Perthame's work and for the MAMBA team, and more generally for the whole domain of mathematics applied to biology.

Marie Doumic was a plenary speaker at the ECMTB 2014 (Göteborg, http://ecmtb2014.org/ 600 participants).

Dirk Drasdo was invited speaker at the Systems Biology of Human Diseases conference (Harvard University, http://www.csb2.org/events/sbhd-2014).

Five articles are noteworthy in terms of bibliometry:

• (*Impact factor 11.2*) F. SCHLIESS, S. HOEHME, S. HENKEL, A. GHALLAB, D. DRIESCH, J. BÖTTGER, R. GUTHKE, M. PFAFF, J. HENGSTLER, R. GEBHARDT, D. HÄUSSINGER, D. DRASDO, S. ZELLMER. Integrated metabolic spatial-temporal model for the prediction of ammonia detoxification during liver damage and regeneration, *Hepatology*, Dec. 2014, vol. 60, no 6, pp. 2040-2051, https://hal.inria.fr/hal-01110646 [17]

• (*Impact factor 10.4*) D. DRASDO, S. HOEHME, J. G. HENGSTLER. How predictive quantitative modeling of tissue organization can inform liver disease pathogenesis, *Journal of Hepatology*, Oct. 2014, vol. 61, no 4, pp. 951-956 [DOI : 10.1016/J.JHEP.2014.06.013], https://hal.inria.fr/hal-01110644 [7]

• (*Impact factor 10.7*) S.R.K. VEDULA, G. PEYRET, I. CHEDDADI, T. CHEN, A. BRUGUÉS, H. HIRATA, H. LOPEZ-MENENDEZ, Y. TOYAMA, L. NEVES DE ALMEIDA, X. TREPAT, C.T. LIM, B. LADOUX. Mechanics of epithelial closure over non-adherent environments, *Nature Communications*, Jan. 2015, vol. 6, art. number 6111[DOI : 10.1038/ncomms7111], http://www.nature.com/ncomms/2015/150122/ncomms7111/abs/ncomms7111.html (open access)

• (*Impact factor 7.5*) L. ROBERT, M. HOFFMANN, N. KRELL, S. AYMERICH, J. ROBERT, M. DOUMIC. Division in Escherichia coli is triggered by a size-sensing rather than a timing mechanism, in "BMC Biology", 2014, vol. 12, no 1, 17 p. [DOI : 10.1186/1741-7007-12-17], https://hal.inria.fr/hal-00981312 [16]

• (*Impact factor 9.3*) R. H. CHISHOLM, T. LORENZI, A. LORZ, A. K. LARSEN, L. ALMEIDA, A. ESCARGUEIL, J. CLAIRAMBAULT. Emergence of drug tolerance in cancer cell populations: an evolutionary outcome of selection, nongenetic instability and stress-induced adaptation, *Cancer Research* (Mathematical oncology), 10p.+suppl. mat., in press, Jan. 2015, https://hal.archives-ouvertes.fr/hal-01111271 [33]

6.2. Cancer

Participants: Luís Lopes Neves de Almeida, José Luis Avila Alonso [DISCO Inria team], Catherine Bonnet [DISCO Inria team], Rebecca Chisholm, Jean Clairambault, François Delhommeau [Hæmatology department, St Antoine Hospital, Paris], Luna Dimitrio [former PhD student and Mamba member], Ján Eliaš, Alexandre Escargueil [Cancer biology and therapeutics lab, St Antoine Hospital, Paris], Pierre Hirsch [Hæmatology department, St Antoine Hospital, Paris], Michal Kowalczyk [Univ. Santiago de Chile], Annette Larsen [Cancer biology and therapeutics lab, St Antoine Hospital, Paris], Tommaso Lorenzi, Alexander Lorz, Anna Marciniak-Czochra [Univ. Heidelberg], Roberto Natalini [IAC-CNR, Univ. Tor Vergata, Rome], Silviu Iulan Niculescu [DISCO Inria team], Hitay Özbay [Bilkent Univ., Ankara], Benoît Perthame, Andrada Maran, Fernando Quirós [Univ. Autónoma de Madrid], Michèle Sabbah [Cancer biology and therapeutics lab, St Antoine Hospital, Paris], Thomas Stiehl [Univ. Heidelberg], Min Tang [Jiaotong University, Shanghai], Emmanuel Trélat [LJLL, UPMC], Nicolas Vauchelet, Romain Yvinec [INRA Tours].

6.2.1. Drug resistance.

We have continued to develop our phenotypically based models of drug-induced drug resistance in cancer cell populations, representing their Darwinian evolution under drug pressure by integro-differential equations. In one of them [40], a 1D space variable has been added to the phenotypic structure variable to account for drug diffusion in tumour spheroids. In another one [33], where deterministic and agent-based modelling are processed in parallel, we have considered a physiologically based 2-dimensional phenotypic structure variable, in order to take account of previously published biological observations on (reversible) drug tolerance persistence in a population of non-small cell lung cancer (NSCLC) cells [0], reproducing the observations and assessing the model by testing biologically based hypotheses. Together with ongoing work with E. Trélat and A. Lorz on drug therapy optimisation, using such phenotype-based models to overcome drug resistance, this has represented a significant part of our work on the subject, which is conducted in close collaboration with the INSERM-UPMC team “Cancer biology and therapeutics” (A. Larsen, A. Escargueil, M. Sabbah) at St Antoine Hospital.

6.2.2. Reversible drug resistance and fractional killing in tumor cell population treatment.

We developed a model of drug resistance in TRAIL (TNF-Related Apoptosis Induced-Ligand) treatment in HeLa cell lines. The TRAIL signal transduction pathway is one of the best studied apoptosis pathways and hence permits detailed comparisons with data. Our model was able to explain experimental observations fractional killing and cell-to-cell variability, and predicted reversible resistance [3]. (Work in close collaboration with G. Batt and S. Stoma from the Inria team LIFEWARE.)

6.2.3. Radiotherapy.

Radiation is still a major treatment in cancer. We explored by extensive computer simulations using an agent-based model the consequences of spatially inhomogeneous irradiation. The model predicted that in the case of different competing sub-populations, namely cancer stem cells with unlimited division capacity, and cancer cells with limited division capacity, inhomogeneous radiation focusing higher doses at the tumour centre and lower doses at the tumour periphery should outperform homogeneous irradiation [12]. Cancer stem cells are believed to have a longer cell cycle duration than cancer cells, and are less radiosensitive than cancer cells, which is why they often survive radiation and lead to tumour relapse.

6.2.4. Intercellular interactions in epithelio-mesenchymal transition (EMT).

A PhD thesis on this subject, co-supervised by L. Almeida and M. Sabbah (INSERM team “Cancer biology and therapeutics”, St Antoine) has begun at Fall. It is also based on phenotype-structured modelling of Darwinian evolution in cancer cell populations.

6.2.5. Interactions between tumour cell populations and their cellular micro-environment.

A phenotype-structured model of the interactions between a brest cancer cell population (MCF7 cultured cells, collaboration with M. Sabbah, St Antoine Hospital) and its adipocyte stroma support cell population has been developed (T. Lorenzi, J. Clairambault), which, beyond submitted proposals (ANR, Emergence Paris-Sorbonne Universités call), will be studied and experimentally identified in a forthcoming internship (January-June 2015) and PhD thesis in applied mathematics.

[0] Sharma *et al.*, Cell, April 2010

6.2.6. Hele-Shaw model of tumour growth.

In the growing field of mathematical analysis of mechanical domain of tumor growth, we focus on the rigorous link between cells models, relying on mechanical properties of cells, and free boundary problem, where the tumor is described by the dynamics of its boundary. The latter model is referred to Hele-Shaw model [44]. Benoît Perthame, Min Tang and Nicolas Vauchelet have proved the rigorous derivation of a geometric model of the Hele-Shaw type for a model with viscoelastic forces, constructing analytically traveling wave solutions of the Hele-Shaw model of tumor growth with nutrient that explain theoretically the numerical results observed. The limiting model exhibits travelling waves, which have been investigated in [43]. Another interesting feature for this model is the transversal instability occurring when the spatial dimension is greater than 1. Together with Fernando Quirós (Univ. Autónoma de Madrid), the aforementioned have also formulated a Hele-Shaw type free-boundary problem for a tumor growing under the combined effects of pressure forces, cell multiplication and active motion, the latter being the novelty of this study [61]. In order to understand the emergence of instabilities in the Hele-Shaw model with nutrients, Michal Kowalczyk (Univ. Chile, Santiago), Benoît Perthame and Nicolas Vauchelet have studied a related model of thermo-reactive diffusion where they can study the spectrum of the linearized system around a traveling wave and in which they can compute the transition to instability in terms of a parameter related to the ratio between heat conduction and molecular diffusion. However, the rigorous study of such instabilities for the whole system of equations is not reachable for the moment; only a study for a simplified model has been performed in [39].

6.2.7. Modelling and control of acute myeloblastic leukæmia (AML).

The collaboration with the Disco project-team has been continued, leading to one book chapter [25], four conference proceedings [21], [22], [23], [24] and JL Avila Alonso's PhD thesis defence.

In more detail:
Starting initially from a PDE model of hematopoiesis designed by Adimy *et al.*[0], we have derived several models of healthy or cancer cell dynamics in hematopoiesis and performed several stability analyses.

We have proposed in [25] a new mathematical model of the cell dynamics in acute myeloid leukæmia (AML) which takes into account the four different phases of the proliferating compartment as well the fast self-renewal phenomenon frequently observed in AML. As was the case in [25] this model is transformed into a distributed delay system and was analyzed here with input-output techniques. Local stability conditions for an equilibrium point of interest are derived in terms of a set of inequalities involving the parameters of the mathematical model.

We have also studied a coupled delay model for healthy and cancer cell dynamics in AML consisting of two stages of maturation for cancer cells and three stages of maturation for healthy cells. For a particular healthy equilibrium point, locally stability conditions involving the parameters of the mathematical model have been obtained [22], [23].

We have performed in [21] a stability analysis of both the PDE model of healthy hæmatopoiesis and a coupled PDE model of healthy and cancer cell dynamics. The stability conditions obtained here in the time domain strengthen the idea that fast self-renewal plays an important role in AML.

A time-domain stability analysis by means of Lyapunov-Krasovskii functionals has been performed on the delay system modeling healthy hematopoiesis for a strictly positive equilibrium point of interest.

Furthermore, a working collaboration on AML modelling with Anna Marciniak-Czochra (Univ. Heidelberg) was also initiated by the end of 2014 by a visit of three of us (C. Bonnet, J. Clairambault, T. Lorenzi) to Heidelberg and a visit of T. Stiehl, A. Marciniak-Czochra PhD student, to Paris. The topics we plan to investigate are, beyond the role of fast self renewal in AML cell populations, the part played by clonal heterogeneity in leukæmic cell populations and the issues it raises in therapeutics, a well known clinical problem in clinical hæmatology.

[0] Adimy, M., Crauste, F., El Abllaoui, A. Discrete maturity-structured model of cell differentiation with applications to acute myelogenous leukemia, *J. Biol. Sys.*, 16(3):395-424, 2008

Let us also mention that on the subject of early leukæmogenesis, Andrada Qillas Maran has undertaken a PhD thesis under the supervision of J. Clairambault and B. Perthame. Models relying on piecewise deterministic Markov processes (PDMPs), designed and studied by R. Yvinec (INRA Tours) for the single-cell part of the model under construction, will be used in collaboration with him. Our clinical referents in hæmatology for this PhD work are F. Delhommeau and P. Hirsch (St Antoine Hospital).

6.2.8. *The p53 protein spatio-temporal dynamics.*

The development of our molecular-based model of the spatio-temporal intracellular dynamics of the p53 protein (the so-called “guardian of the genome") has been continued [55], [9], leading us also, more generally, to propose a modelling frame dedicated to the dynamics of intracellular proteins and their gene regulatory networks [8].

6.2.9. *Others.*

In a collaboration with ANGE, B. Perthame has studied a data assimilation algorithm for multidimensional hyperbolic conservation laws using kinetic schemes and kinetic formulations.

6.3. Aggregation kinetics

Participants: Tom Banks, Thibault Bourgeron, Marc Hoffmann, Marie Doumic-Jauffret, Nathalie Krell, Benoît Perthame, Stéphanie Prigent, Human Rezaei, Nathalie Robert, Léon Matar Tine [Univ. Lyon and Dracula Inria team], Jorge Zubelli [IMPA, Rio de Janeiro].

6.3.1. *Time Asymptotics for Fragmentation Equations*

Fragmentation and growth-fragmentation equations is a family of problems with varied and wide applications. This paper is devoted to description of the long time time asymptotics of two critical cases of these equations, when the division rate is constant and the growth rate is linear or zero. The study of these cases may be reduced to the study of the following fragmentation equation:

$$\frac{\partial}{\partial t}u(t,x)+u(t,x)=\int_x^\infty k_0(xy)u(t,y)dy.$$

Using the Mellin transform of the equation, we determine the long time behavior of the solutions. Our results show in particular the strong dependence of this asymptotic behavior with respect to the initial data.

6.3.2. *Estimating the division rate in a size-structured population.*

The problem which was considered in [5] consists in estimating the division rate from the stable size distribution of the population, which is easily measured, but non-smooth. We propose a method based on the Mellin transform for growth-fragmentation equations with self-similar kernels. We build a sequence of functions which converges to the density of the population in division, simultaneously in several weighted L^2 spaces, as the measurement error goes to 0. This improves previous results for self-similar kernels [0] and allows us to understand the partial results for general fragmentation kernels [0]. Numerical simulations confirm the theoretical results. Moreover, our numerical method is tested on real biological data, arising from a bacteria growth and fission experiment.

6.3.3. *What governs bacterial growth? The “sizer” vs the “timer” model*

We applied the previously seen inverse problem methodology [5] to a fundamental biological problem: what governs the bacterial growth?

[0] Perthame and Zubelli, *Inv. Prob.*, 2007
[0] Doumic and Tine, *J. Math. Biol.*, 2012

Many organisms coordinate cell growth and division through size control mechanisms: cells must reach a critical size to trigger a cell cycle event. Bacterial division is often assumed to be controlled in this way, but experimental evidence to support this assumption is still lacking. Theoretical arguments show that size control is required to maintain size homeostasis in the case of exponential growth of individual cells. Nevertheless, if the growth law deviates slightly from exponential for very small cells, homeostasis can be maintained with a simple ‘timer’ triggering division. Therefore, deciding whether division control in bacteria relies on a ‘timer’ or ‘sizer’ mechanism requires quantitative comparisons between models and data.

The timer and sizer hypotheses find a natural expression in models based on partial differential equations. Here we test these models with recent data on single-cell growth of Escherichia coli. We demonstrate that a size-independent timer mechanism for division control, though theoretically possible, is quantitatively incompatible with the data and extremely sensitive to slight variations in the growth law. In contrast, a sizer model is robust and fits the data well. In addition, we tested the effect of variability in individual growth rates and noise in septum positioning and found that size control is robust to this phenotypic noise.

Confrontations between cell cycle models and data usually suffer from a lack of high-quality data and suitable statistical estimation techniques. In the study [16] we had overcome these limitations by using high precision measurements of tens of thousands of single bacterial cells combined with recent statistical inference methods to estimate the division rate within the models. We therefore provided the first precise quantitative assessment of different cell cycle models.

6.3.4. Size distribution of amyloid fibrils. Mathematical models and experimental data.

More than twenty types of proteins can adopt misfolded conformations, which can co-aggregate into amyloid fibrils, and are related to pathologies such as Alzheimer’s disease. In [15], we surveyed mathematical models for aggregation chain reactions, and discussed the ability to use them to understand amyloid distributions. Numerous reactions have been proposed to play a role in their aggregation kinetics, though the relative importance of each reaction in vivo is unclear: these include activation steps, with nucleation compared to initiation, disaggregation steps, with depolymerization compared to fragmentation, and additional processes such as filament coalescence or secondary nucleation. We have statistically analysed the shape of the size distribution of prion fibrils, with the specific example of truncated data due to the experimental technique (electron microscopy). A model of polymerization and depolymerization succeeds in explaining this distribution. It is a very plausible scheme though, as evidenced in the review of other mathematical models, other types of reactions could also give rise to the same type of distributions.

To clarify how these fibrils are able to incorporate additional units, prion fibril aggregation and disaggregation kinetics were experimentally studied using Static Light Scattering (SLS) [45]. Values that are functions of $\sum i^2 c_i$ (for $i > 0$) with c_i being the concentration of fibrils of size i, were then measured as a function of time. An initial model, adapted from the Becker-Döring system that considers all fibrils to react similarly is not able to reproduce the observed in vitro behaviour. Our second model involves an additional compartment of fibrils unable to incorporate more prion units. This model leads to kinetic coefficients which are biologically plausible and correctly simulates the first experimental steps for prion aggregation.

In the formation of large clusters out of small particles, the initializing step is called the nucleation, and consists in the spontaneous reaction of agents which aggregate into small and stable polymers called nucleus. After this early step, the polymers are involved into a bunch of reactions such as polymerization, fragmentation and coalescence. Since there may be several orders of magnitude between the size of a particle and the size of an aggregate, building efficient numerical schemes to capture accurately the kinetics of the reaction is a delicate step of key importance. In [29], we propose a conservative scheme, based on finite volume methods on an adaptive grid, which is able to render out the early steps of the reaction as well as the later chain reactions.

6.4. Liver organ modelling

Participants: Noémie Boissier, Dirk Drasdo, Géraldine Cellière, Adrian Friebel, Group Heinzle [Univ. Saarbruecken, Germany], Group Hengstler [IfADo, Germany], Stefan Hoehme, Tim Johann, Group Klingmueller [German Cancer Center, Heidelberg], Johannes Neitsch, Group Reo [Inria Paris - Rocquencourt], Paul Van Liedekerke, Eric Vibert [Hopital Paul Brousse], Yi Yin, Group Zerial [Max-Planck Inst. for Molecular Genetics, Dresden, Germany], Groups Iflow, Notox, Vln.

6.4.1. Ammonia detoxification after drug-induced damage.

The model for ammonia detoxification after drug-induced damage (see above) identified a systematic deviation between data and results that would be expected from the current standard model for ammonia detoxification in healthy liver [0] ([17], [6]) (see also comments/editorials in [0]). The findings triggered a series of new experiments identifying reversibility of the glutamate-dehydrogenase reaction in hepatocytes, and in blood (Ghallab et. al., subm.). It could be shown in an animal model that the newly recognized reactions can be therapeutically used to significantly reduce the concentration of toxic ammonia after drug-induced damage. (Work in close collaboration with partners of the project VLN (BMBF, Germany) and EU-NOTOX.

6.4.2. Systematic analysis strategies permitting quantitative conclusions in systems medicine and biology.

Based on the examples from liver regeneration after drug-induced damage [57] [17]) systematic iterative strategies can be inferred to enable identification of mechanisms underlying complex processes in spatial temporal tissue organisation and organ functioning. These use an iterative application of a pipeline of imaging, image analysis and modeling, quantitative models by parameterization of model components by measurable parameters for which the physiological ranges are known, and systematic simulated parameter sensitivity analyses [7].

[0] Haeussinger D., Eur. J. Biochem, 1983; Gebhardt R and Mecke, D. EMBO J 1983
[0] Wierling, C. Hepatology, 60(6) 2014; and: Widera, A., EXCLI Journal, 13, 2014

MANAO Project-Team

5. New Results

5.1. Highlights of the Year

We are still developping our expertise in fitting techniques. As an illustration, we have solved of a long-standing problem in fluid capture: the non-invasive three-dimensional digitization of dynamic gas flows *including their three-dimensional velocity fields* [17] (cf. Figure 8). We solve the three-dimensional flow tracking problem by fitting a full 3D Navier-Stokes simulation to the acquired data. To our knowledge, this is a world-first in this area that considerably improves the results by incorporating high-level prior knowledge into the estimate. The resulting mathematical framework can be generalized easily and lends itself to editing operations. The technique has applications, e.g., in aerospace engineering. We are exploring the possibilities with ONERA, the French space agency. In fact, parts of the developed techniques have been validated by them and are now being installed in a wind tunnel facility for real-world tests.

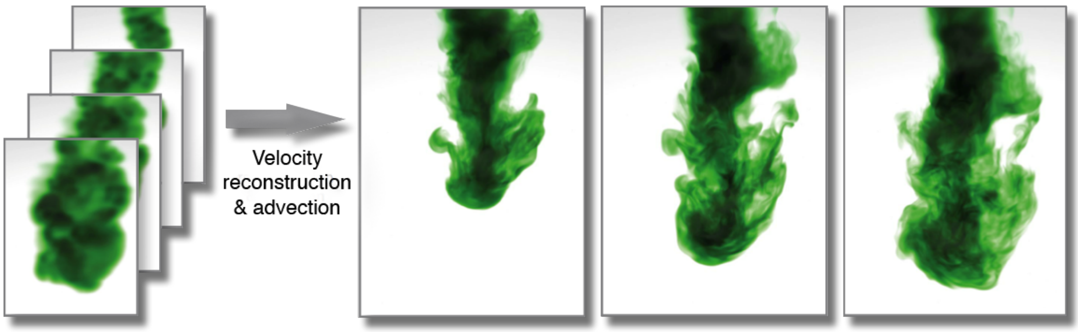


Figure 8. Low-resolution captures obtained by tomographic scanning (left) are used as inputs to our method which estimates physically plausible dense velocity fields. Such velocity fields fully determine the fluid state and can be applied in a variety of applications including fluid super-resolution (right) allowing capture to be integrated into pipelines for visual effects simulation.

This year, the collaboration between Optics and Computer Graphics has grown to a now long-term project, under the initiative of the MANAO team. First, from an institutional point of view, a framework agreement has been signed the 10th of July 2014 between the IOGS and Inria. This is an important and institutional recognition of the potential trans-disciplinary impacts of our work. Second, we have begun to set-up the COEL experimentation facility inside the LP2N laboratory. It has been made possible thanks to the support of the "Région Aquitaine" and upcoming supports from l'"Initiative d'excellence de l'université de Bordeaux". With this trans-disciplinary experimentation facility – rather unique in Europe – we can now put into practice a long-term vision of the researches that we want to achieve.

In term of visibility, we managed to published our first paper in the Optics scientific community [15], highlighting our trans-disciplinary research. We have also been part of the final and transnational exhibition of the V-Must.net network of excellence: Keys2Rome - http://keys2rome.eu. It was launched simultaneously in Rome, Sarajevo, Amsterdam and Alexandria on September 23, 2014. The exhibition uses immersive technology to present and connect these regional cultures within the Roman Empire, highlighting their diversity and commonality over centuries of Roman rule. Our spatial augmented reality solution [21] was included in this event.

5.2. Analysis and Simulation

5.2.1. Importance Sampling of Realistic Light Sources

Realistic images can be rendered by simulating light transport with Monte Carlo methods. The possibility to use realistic light sources for synthesizing images greatly contributes to their physical realism. Among existing models, the ones based on environment maps and light fields are attractive due to their ability to capture faithfully the far-field and near-field effects as well as their possibility of being acquired directly. Since acquired light sources have arbitrary frequencies and possibly high dimensions (4D), using such light sources for realistic rendering leads to performance problems. We have investigated [12] how to balance the accuracy of the representation and the efficiency of the simulation (cf. Figure 9). The work relies on generating high quality samples from the input light sources for unbiased Monte Carlo estimation [74]. This is a foundation work that has leaded to new sampling techniques for physically-based rendering with time-varying environment lighting [73] and light field light sources. The results show that physically accurate rendering with realistic light sources can be achieved in real time.

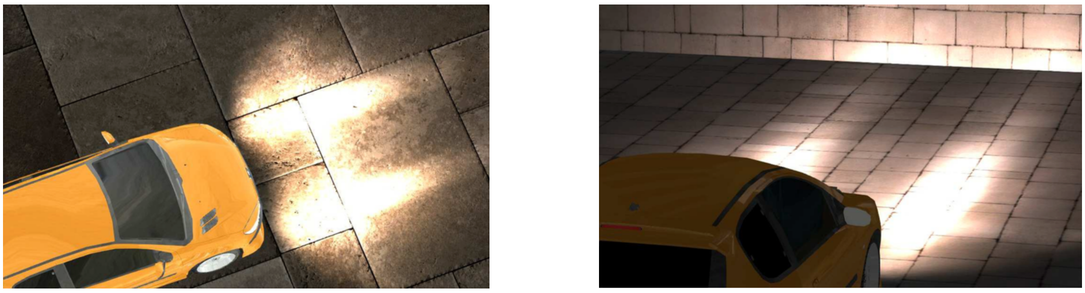


Figure 9. Our new light importance sampling technique estimates direct lighting interactively (7-9 fps) with only 200 samples per pixel that are distributed among the different images of the light field luminaire. The car headlights are represented by the same light field composed of 11 $\times$ 9 images (256 $\times$ 256 pixels).

5.2.2. Frequency Analysis of Light Scattering and Absorption

We have proposed [14] an innovative analysis of absorption and scattering of local light fields in the Fourier domain, and derived the corresponding set of operators on the covariance matrix of the power spectrum of the light field. This analysis brings an efficient prediction tool for the behavior of light along a light path in participating media. We leverage this analysis to derive proper frequency prediction metrics in 3D by combining per-light path information in the volume. Our key contribution is to show that analyzing local light fields in the Fourier domain reveals the consistency of illumination in such media, and provides a set of simple and useful rules to be used to accelerate existing global illumination methods.

5.3. Acquisition and Display

5.3.1. Three-Dimensional, Dynamic, Full State Fluid Capture and Manipulation

Participant: I. Ihrke

We have explored [17] the connection between fluid capture, simulation and proximal methods, a class of algorithms commonly used for inverse problems in image processing and computer vision. Our key finding is that the proximal operator constraining fluid velocities to be divergence-free is directly equivalent to the pressure-projection methods commonly used in incompressible flow solvers. This observation lets us treat the inverse problem of fluid tracking as a constrained flow problem all while working in an efficient, modular framework. In addition it lets us tightly couple fluid simulation into flow tracking, providing a global prior that significantly increases tracking accuracy and temporal coherence as compared to previous techniques. We demonstrate how we can use these improved results for a variety of applications, such as re-simulation, detail enhancement, and domain modification. We furthermore give an outlook of the applications beyond fluid tracking that our proximal operator framework could enable by exploring the connection of deblurring and fluid guiding.

5.3.2. Measurements and Analysis of Retro-reflective Materials

Participants: L. Belcour, R. Pacanowski

We have compared [15] performance of various analytical retro-reflecting BRDF models to assess how they reproduce accurately measured data of retro-reflecting materials. We have also introduced a new parametrization, the back vector parametrization, to analyze retro-reflecting data and we have shown that this parametrization better preserves the isotropy of data. Furthermore, we have updated existing BRDF models to improve the representation of retro-reflective data. This work was supported by the development of the ALTA library [23].

5.3.3. Kaleidoscopic Imaging

Participants: I. Reshetouski, I. Ihrke

Kaleidoscopes have a great potential in computational photography as a tool for redistributing light rays. In time-of-flight imaging the concept of the kaleidoscope is also useful when dealing with the reconstruction of the geometry that causes multiple reflections. Our work [13] is a step towards opening new possibilities for the use of mirror systems as well as towards making their use more practical. The focus of this work is the analysis of planar kaleidoscope systems to enable their practical applicability in 3D imaging tasks. We have analyzed important practical properties of mirror systems and developed a theoretical toolbox for dealing with planar kaleidoscopes. Based on this theoretical toolbox, we have explored the use of planar kaleidoscopes for multi-view imaging and for the acquisition of 3D objects [90]. The knowledge of the mirrors positions is crucial for these multi-view applications. On the other hand, the reconstruction of the geometry of a mirror room from time-of-flight measurements is also an important problem. We therefore employ the developed tools for solving this problem using multiple observations of a single scene point.

5.3.4. Interactive Spatial Augmented Reality

Participants: B. Ridel, P. Reuter, X. Granier

We have proposed the *Revealing Flashlight* [21], a new 6-degree-of-freedom interaction and visualization technique in spatial augmented reality that helps to reveal the details of cultural heritage artifacts. We locally and interactively highlight them by projecting an expressive visualization. The Revealing Flashlight can be used by archaeologists, for example, to help decipher inscriptions in eroded stones, or by museums (cf. Figure 10) to let visitors interactively discover the features and meta-information of cultural artifacts. A permanent exhibition is now running at the Allard Pierson Museum, and others museums are asking us to set-up similar installations. It was part of the final trans-European showcase of the V-MusT.net project.

5.4. Rendering, Visualization & Illustration

5.4.1. Computing Smooth Surface Contours with Accurate Topology

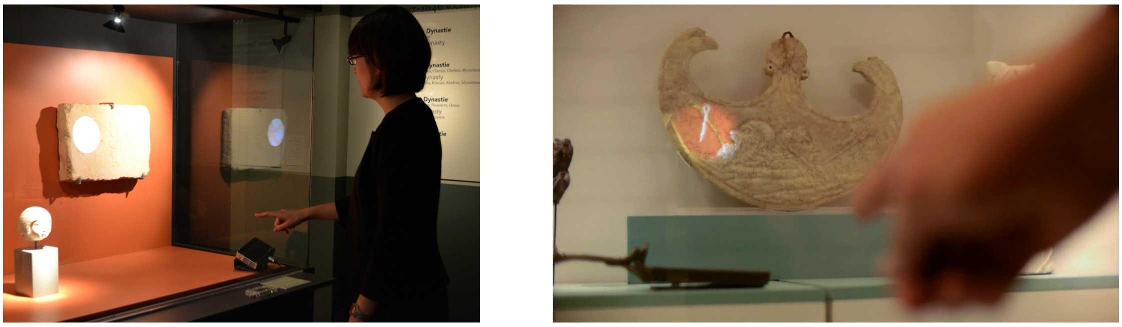


Figure 10. "The Revealing Flashlight" lets visitors explore ancient artifacts interactively. (Left) Allard Pierson Museum - Amsterdam. (Right) Keys2Rome exhibition in Museo dei Fori Imperiali - Roma.

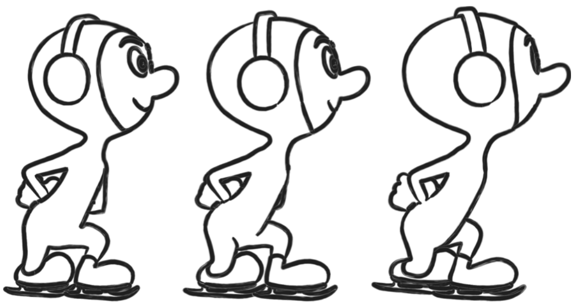


Figure 11. Contours stylized with tapered strokes [16]. Our method avoids classical breaks and gaps, producing more coherent animated strokes. Red © Pixar

We have introduced [16] a method for accurately computing the visible contours of a smooth 3D surface for stylization. This is a surprisingly difficult problem, and previous methods are prone to topological errors, such as gaps in the outline. Our approach is to generate, for each viewpoint, a new triangle mesh with contours that are topologically-equivalent and geometrically close to those of the original smooth surface. The contours of the mesh can then be rendered with exact visibility. The core of the approach is Contour-Consistency, a way to prove topological equivalence between the contours of two surfaces. Producing a surface tessellation that satisfies this property is itself challenging; to this end, we introduce a type of triangle that ensures consistency at the contour. We then introduce an iterative mesh generation procedure, based on these ideas. This procedure does not fully guarantee consistency, but errors are not noticeable in our experiments. Our algorithm can operate on any smooth input surface representation; we use Catmull-Clark subdivision surfaces in our implementation.

5.5. Editing and Modeling

5.5.1. *Tomography-Based Volume Painting*

Participant: I. Ihrke

Although volumetric phenomena are important for realistic rendering and can even be a crucial component in the image, the artistic control of the volume's appearance is challenging. Appropriate tools to edit volume properties are missing, which can make it necessary to use simulation results directly. Alternatively, high-level modifications that are rarely intuitive, e.g., the tweaking of noise function parameters, can be utilized. We have introduced [18] a solution to stylize single-scattering volumetric effects in static volumes. Hereby, an artistic and intuitive control of emission, scattering and extinction becomes possible, while ensuring a smooth and coherent appearance when changing the viewpoint. Our method is based on tomographic reconstruction, which we link to the volumetric rendering equation. It analyzes a number of target views provided by the artist and adapts the volume properties to match the appearance for the given perspectives. Additionally, we describe how we can optimize for the environmental lighting to match a desired scene appearance, while keeping volume properties constant. Finally, both techniques can be combined. We demonstrate several use cases of our approach and illustrate its effectiveness.

5.5.2. *Implicit Skinning*

Participant: G. Guennebaud

In collaboration with IRIT (Toulouse), we extended our *implicit skinning* method to a new approach for interactive character skinning called *elastic implicit skinning*. The method simulates skin contacts between limbs as well as the effect of skin elasticity (Figure 12). In addition, we go a step further towards the automation of the rigging process: our method doesn't require the definition of skinning weights. Elastic implicit skinning takes the best features of the recent implicit skinning method, and makes it robust to extreme character movements. While keeping the idea of implicit skinning, namely approximate the character by 3D scalar fields in which mesh-vertices are appropriately re-projected, we depart from the processing pipeline used so far. Implicit skinning is history independent and uses an initial skinning solution (e.g., linear blending or dual quaternions) to correct vertex positions at each frame. Our new approach is history dependent; the mesh directly tracks the iso-surfaces of the scalar field over time. Technically our solutions include: new implicit surface composition operators and a tangential relaxation scheme derived from the as-rigid-as possible energy. This work [101] has been presented at SIGGRAPH Asia this year.

5.5.3. *Multi-scale Editing*

Participant: G. Guennebaud

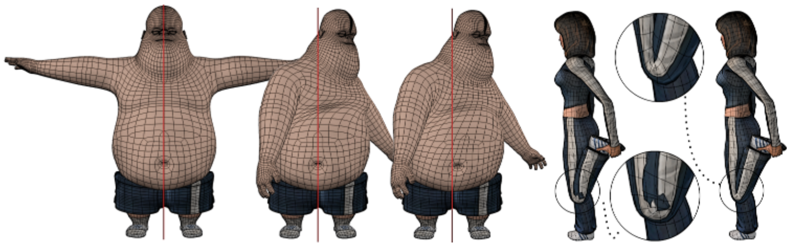


Figure 12. Illustration of the benefits of our novel elastic implicit skinning technique.

In the continuation of our Growing Least Square approach [5] for the multi-scale analysis of shape, we developed a novel tool that enables the direct editing of surface features in large point-clouds or meshes [19]. This is made possible by a novel multi-scale analysis of unstructured point-clouds that automatically extracts the number of relevant features together with their respective scale all over the surface. Then, combining this ingredient with an adequate multi-scale decomposition allows us to directly enhance or reduce each feature in an independent manner. Our feature extraction is based on the analysis of the scale-variations of locally fitted surface primitives combined with unsupervised learning techniques. Our tool may be applied either globally or locally, and millions of points are handled in real-time. The resulting system enables users to accurately edit complex geometries with minimal interaction.

5.5.4. *Manipulation of Anisotropic Highlights*

Participants: B. Raymond, P. Barla, G. Guennebaud, X. Granier

We have developed [20] a system for the direct editing of highlights produced by anisotropic BRDFs, which we call anisotropic highlights. We first provide a comprehensive analysis of the link between the direction of anisotropy and the shape of highlight curves for arbitrary object surfaces. The gained insights provide the required ingredients to infer BRDF orientations from a prescribed highlight tangent field. This amounts to a non-linear optimization problem, which is solved at interactive framerates during manipulation. Taking inspiration from sculpting software, we provide tools that give the impression of manipulating highlight curves while actually modifying their tangents. Our solver produces desired highlight shapes for a host of lighting environments and anisotropic BRDFs.

MARELLE Project-Team

6. New Results

6.1. Highlights of the Year

In June 2014, Yves Bertot received the ACM Software System award, as one of the main contributors to the Coq System, along with Gérard Huet, Thierry Coquand, Christine Paulin-Mohring, Bruno Barras, Jean-Christophe Filliâtre, Hugo Herbelin, Chet. Murthy, and Pierre Castéran.

6.2. Proof and computation

Participants: Laurent Théry [correspondant], Benjamin Grégoire.

We have been continuing our effort to improve the computing power of Coq. This has led to two "computational proof":

The Erdös conjecture for n = 2 was proved this year using a SAT solver. We succeeded to formally prove this instance in Coq independently checking the 3Gb trace of the SAT solver.

The weak Goldbach conjecture was proved last year by Harald Helfgott. This proof requires a computation that the conjecture holds for numbers less than 10^{28}. This is done in two stages. The first one is to verify Goldbach conjecture for numbers less than 10^{18}. The second one is to verify the weak Goldbach conjecture for numbers less than 10^{28} using a ladder with intervals 10^{18}. The second stage has been completely verified in Coq. We are currently working on improving the computation power of Coq to make it possible to perform the first stage in reasonable time.

6.3. Formal verification of automated proof algorithms

Participant: Laurent Théry [correspondant].

We have been interested in proving that the classic 2-Sat problem can be solved in linear time. This leads to proving two classic algorithms:

1. A version of Kosaraju's algorithm that computes the strongly connected components of a directed graph [21],
2. A more direct algorithm that solves the 2-Sat problem that is using unit propagation, proposed by Alvaro del Val [20].

6.4. Formal study of cryptography

Participants: Gilles Barthe [IMDEA], Sonia Belaid [THALES and ENS], François Dupressoir [IMDEA], Pierre-Alain Fouque [Université de Rennes 1 and Institut universitaire de France], Cédric Fournet [Microsoft Research], Benjamin Grégoire [correspondant], Benedikt Schmidt [IMDEA], Pierre-Yves Strub [IMDEA], Nikhil Swamy [Microsoft Research], Mehdi Tibouchi [NTT Secure Platform Laboratories], Santiago Zanella-Béguelin [Microsoft Research], Jean-Christophe Zapalowicz [Inria].

The goal of this work is to provide a friendly tool easily usable by cryptographers without knowledge of formal proof assistants. The idea is to use the techniques formally proved in Certycrypt and to call SMT-provers. We provide two differents tools EasyCrypt and ZooCrypt.

This year, we worked on the following topics:

- Relational program logics, as used in EasyCrypt, have been used for mechanizing formal proofs of various cryptographic constructions. In [15], we present rF*, a relational extension of F*, a general-purpose higher-order stateful programming language with a verification system based on refinement types. The distinguishing feature of rF* is a relational Hoare logic for a higher-order, stateful, probabilistic language.
- Fault Attacks are attacks in which an adversary with physical access to a cryptographic device, say a smartcard, tampers with the execution of an algorithm to retrieve secret material. In [13] we propose a new approach for finding fault attacks based on fault conditions. Using the method, we discover multiple fault attacks on RSA and ECDSA. Several of the attacks found by our tool are new. In [14], we propose a new counter measure to make RSA-PSS provably secure against non-random faults. We also prove the result using EasyCrypt.
- Many algorithms, particularly in cryptography, admit very efficient batch versions that compute simultaneously the output of the algoritms on a set of inputs. AutoBatch is a tool that computes highly optimized batch verification algorithms for pairing based signature schemes. In [12], we use EasyCrypt to formalise the methods used by AutoBatch and to automatically certify the result of the transformation performed by AutoBatch.
- We study the problem of automatically verifying higher-order masking countermeasures which is used to protect implementations where the attacker can observe intermediate computations (like in a smartcard). We propose an efficient method to check the correctness and the security of masked implementation. This work has been submitted to EuroCrypt 2015. We start the ANR BRUTUS on this subject.

6.5. Formalization of Bourbaki's sets and ordinals

Participant: José Grimm.

In previous years we developed a formal library describing the parts of the Bourbaki books on set theory, cardinals and ordinals. We completed it by adding the definition of real numbers using Dedekind cuts. The important properties we showed that $\mathbf{R}$ is an ordered Archimedean field, that every non-empty bounded subset has a least upper bound, that every Cauchy sequence has a limit, and that the intermediate value theorem holds.

It follows that every positive real number has positive square root. We give a pair of adjacent sequences that converges to this square root. For instance $\sqrt{2}$ is irrational, and we get a pair of rational adjacent sequences that converges to it. This produces an explicit order isomorphism $\mathbf{Q}^* \to \mathbf{Q}$. The number of such isomorphisms is equal to the power of the continuum (the cardinal of $\mathbf{R}$) [18].

6.6. Stern-Brocot and Fibonacci sequences

Participant: José Grimm.

We constructed an explicit bijection $\mathbf{N} \to \mathbf{Q}$, first in the framework of the Bourbaki project (see above), then in Ssreflect. Every positive rational number x can uniquely be written as a quotient s_n/s_{n+1}. This result was established by Dijkstra who stated it in an obfuscated way. It was shown years before by Stern. It is possible to compute s_n/s_{n+1} without computing numerator and denominator separately, by considering the sequences of bits of n from left to right or from right to left. Truncating the binary expansion of n yields a sequence of approximations to s_n/s_{n+1} (this was studied by Brocot, and the so-called Stern-Brocot tree is an alternative representation of rational numbers). We implemented the work of Dijkstra and Stern in Coq [17].

We also studied how a number can be represented by a sequence of other numbers (for instance as a sum of distinct Fibonacci numbers, with or without constraints). The number of ways of writing n as a sum of powers of two, each power of two being used at most twice, is s_{n+1}. These results are presented in [17].

6.7. Formal proof that e and π are transcendental

Participants: Sophie Bernard, Laurence Rideau.

We constructed formal proofs that π is irrational, e is transcendental, and π is transcendental. These proofs share a common initial pattern, where rationality or algebraicity of the mathematical constants are shown to imply the existence of a sequence of positive integers that must decrease indefinitely.

This proof development is an opportunity to study the interplay between several existing libraries about algebraic structures and analysis: the ssreflect library for algebra and the Coquelicot library for calculus. Moreover, the proof that π is transcendental was an occasion to test the newly developed module on symmetric polynomials by P.-Y. Strub at IMDEA.

6.8. Fast computation of π

Participant: Yves Bertot.

In the previous year, we studied a proof that π could be approximated with a fast converging sequence based on arithmetic geometric means. This year we described a proof that rounding errors during this computation could be guaranteed as small as needed, based on a study of derivatives. This approach provides a fruitful alternative to interval-based approaches. The result was published in [16].

We also completed a journal paper on various ways to observe and compute the number π [7].

6.9. Decision procedures for polynomials

Participant: Yves Bertot.

Following up on the work in previous years around Bernstein Polynomials, we implemented a decision procedure for guaranteeing the sign of a polynomial function inside an interval, using Bernstein polynomials and dichotomy. In the long run, we hope to explore two approaches, one based on the off-line computation of certificates for sub-intervals (these certificates are easy to verify), and one based on implementing computational reflection. This approach should also generalize quite easily to multi-variate polynomials.

MASAIE Project-Team

5. New Results

5.1. Highlights of the Year

The estimation of sequestered parasite population has been a challenge for the biologist and modeler, with many authors having studied this problem. The difficulty is that the infected erythrocyte leaves the circulating peripheral blood and binds to the endothelium in the microvasculature of various organs. A measurement of Plasmodium falciparum parasitaemia taken from a blood smear therefore samples young parasites only and there is no clinical methods to measure the sequestered parasites. We have developed a simple tool to estimate the sequestered parasites and hence the total parasite burden for *Plasmodium falciparum* malaria patients. We have also given a method to estimate a crucial parameter in the model of infection. This parameter β can be thought as the “transmission/invading" factor between merozoites and erythrocytes. This work [9] has been published in "Mathematical Biosciences and Engineering".

5.2. Modeling the use of Wolbachia for controlling the incidence of dengue

We continued research on modeling the introduction of *Wolbachia* in a population of *Aedes Aegypti*. This research is done in collaboration with FGV (Fundação Getulio Vargas), Fiocruz (Fondation Oswaldo Cruz) and UFF (Universidade Federal Fluminense) in Rio de Janeiro (Brazil) [16].

Wolbachia is a bacteria which infects arthropod species, including a high proportion of insects (60% of species). Its interactions with its hosts are often complex, and in some cases it is considered as an endosymbiont. The unique biology of *Wolbachia* has attracted a growing number of researchers interested in questions ranging from the evolutionary implications of infection through to the use of this agent for pest and disease control: a public web site has been funded by the National Science Foundation of Australia, and a research in pubmeb (http://www.ncbi.nlm.nih.gov/pubmed) typing `wolbachia` gives 1889 results.

While *Wolbachia* is commonly found in many mosquitoes it is absent from the species that are considered to be of major importance for the transmission of human pathogens. The successful introduction of a life-shortening strain of *Wolbachia* into the dengue vector *Aedes aegypti* that decreases adult mean life has recently been reported.

Moreover it is estimated that the population of mosquitoes harboring *Wolbachia* is less efficient to transmit dengue [18], [21], [22], [25]. Then it is considered that using *Wolbachia* can be a viable option for controlling the incidence of the dengue.

We consider an alternative infection (by Wolbachia) model which exhibits monotonous properties. This model is designed to take into account both the biology of this infection and any available data. The objective is to use this model for predicting the sustainable introduction of this bacteria. We provide a complete mathematical analysis of the model proposed and give the basic reproduction ratio $\mathcal{R}_0$ for *Wolbachia*. We observe a bistability phenomenon. Two equilibria are asymptotically stable: an equilibrium where all the population is uninfected and an equilibrium where all the population is infected. A third unstable equilibrium exists. We provide a lower bound for the basin of attraction of the desired infected equilibrium. We are in a backward bifurcation situation. The bistable situation occurs with natural biological values for the parameters. Our model is an example of an epidemiological model with only vertical transmission.

This infection model is then connected with a classical dengue model. We prove that for the complete model the equilibrium with *Wolbachia* for the mosquitoes and without dengue for the human is asymptotically stable. We prove that, if a sufficiently large population of infected (by Wolbachia) mosquitoes is introduced, dengue will disappear.

These results have been obtained in collaboration with Pierre-Alexandre Bliman (FGV, Inria); Moacyr Silva (FGV), Claudia Codeço (Fiocruz), Max Souza (UFF) and Jair Koiller (FGV).

5.3. Estimating the proportion of susceptible individuals for a dengue epidemic

Starting from the multi-scaled dengue system, we construct a pair of observers to estimate the dynamics of the disease. The nature of both the observers and the multi-scaled system allows to estimate both the number of susceptible and recovered hosts, as well as to provide information on the vector population, using only infected population data. Numerical simulations have been used to illustrate the performance of the observers.

5.4. Singular value decomposition in dynamic epidemiology: arboviral diseases with human circulation

We introduce a matrix that combines information about human circulation and the epidemiological situation at the nodes of a metapopulational model for an arboviral disease. Its singular value decomposition allows relationships between three basic reproduction numbers $\mathcal{R}_0$: local(s), uniform, and network. The onset of an arboviral disease is strongly dependent on the network characteristics. We present a naive "early warning" criterion for the outbreak at a given node, aiming to promote a discussion on the role of left and right singular vectors. This work is done by the Brazil-France Capes/Cofecub team.

5.5. Analysis of a schistosomiasis infection model

The global mathematical analysis of a schistosomiasis infection model that involves human and intermediate snail hosts as well as an additional mammalian host and a competitor snail species has been done by constructing Lyapunov functions and using properties of K monotone systems. We derived the basic reproduction number $\mathcal{R}_0$ for the deterministic model, and establish that the global dynamics are completely determined by the values of $\mathcal{R}_0$. This mathematical analysis of the model gives insight about the epidemiological consequences of the introduction of a competitor resistant snail species. We gave the characteristics of the competitor resistant snail species that can be used to eliminate the disease [11].

5.6. Multi-stages and multi-strains epidemic models

The model SI (Susceptible-Infected) is one of the most important and used epidemiological models. We gave a complete analysis of the stability of the model with a non-linear incidence and two classes of infected individuals [12].

We have also studied SIS, SIR and MSIR models with bilinear incidence and varying population, with n different pathogen strains of an infectious disease, with or without vertical transmission. For these classes of models, we have proved that under generic conditions a competitive exclusion principle holds. To each strain a basic reproduction ratio can be associated. It corresponds to the case where only this strain exists. The basic reproduction ratio of the complete system is the maximum of the individual basic reproduction ratios. Actually we have also defined an equivalent threshold for each strain. The winner of the competition is the strain with the maximum threshold. It turns out that this strain is the most virulent, i.e., this is the strain for which the endemic equilibrium gives the minimum population for the susceptible host population. This can be interpreted as a pessimization principle [10].

A mathematical multi-patches model for highland malaria in Kenya has been developed and analysed in [13] and [14].

MATHERIALS Team

5. New Results

5.1. Electronic structure calculations

Participants: Eric Cancès, Virginie Ehrlacher, David Gontier, Claude Le Bris, Gabriel Stoltz.

In electronic structure calculation as in most of our scientific endeavours, we pursue a twofold goal: placing the models on a sound mathematical grounding, and improving the numerical approaches.

E. Cancès and N. Mourad have mathematically analyzed the density functional perturbation theory, both in the non-degenerate case (that is, when the Fermi level is not an eigenvalue of the Kohn-Sham hamiltonian) and in the degenerate case. They have in particular proved that Wigner's 2n+1 rule holds in both cases. D. Gontier has obtained a complete, explicit, characterization of the set of spin-polarized densities for finite molecular systems. This problem was left open in the pioneering work of von Barth and Hedin setting up the Kohn-Sham density functional theory for magnetic compounds. He has also extended a previous work by Anantharaman and Cancès, and proved the existence of minimizers for the spin-polarized Kohn-Sham model in the presence of a magnetic field within the local spin density approximation.

E. Cancès has pursued his long-term collaboration with Y. Maday (UPMC) on the numerical analysis of electronic structure models. With L. He (ENPC) and R. Chakir (IFSTTAR), they have designed and analyzed a two-grid methods for nonlinear elliptic eigenvalue problems, which can be applied, in particular, to the Kohn-Sham model. Some numerical tests demonstrating the interest of the approach have been performed with the Abinit software. Together with G. Dusson (UMPC), B. Stamm (UMPC), and M. Vohralík (Inria), they have designed a new post processing method for planewave discretizations of nonlinear Schrödinger equations, and used it to compute sharp *a posteriori* error estimators for both the discretization error and the algorithmic error (convergence threshold in the iterations on the nonlinearity).

Implicit solvation models aims at computing the properties of a molecule in solution (most chemical reactions take place in the liquid phase) by replacing all the solvent molecules but the few ones strongly interacting with the solute, by an effective continuous medium accounting for long-range electrostatics. E. Cancès, Y. Maday (Paris 6), and B. Stamm (Paris 6) have recently introduced a very efficient domain decomposition method for the simulation of large molecules in the framework of the so-called COSMO implicit solvation models. In collaboration with F. Lipparini (UPMC, B. Mennucci (Department of Chemistry, University of Pisa) and J.-P. Picquemal (Paris 6), they have implemented this algorithm in widely used computational software products (Gaussian and Tinker). The extension of this method to other implicit solvation models is work in progress.

Claude Le Bris, in collaboration with Pierre Rouchon (Ecole des Mines de Paris), has pursued the study of a new efficient numerical approach, based on a model reduction technique, to simulate high dimensional Lindblad type equations at play in the modelling of open quantum systems. The specific case under consideration is that of oscillation revivals of a set of atoms interacting resonantly with a slightly damped coherent quantized field of photons. The approach may be employed for other similar equations. Current work is directed towards other numerical challenges for this type of problems.

5.2. Computational Statistical Physics

Participants: Thomas Hudson, Frédéric Legoll, Tony Lelièvre, Mathias Rousset, Gabriel Stoltz.

The work of the team in this area is concentrated on two new directions: the sampling of reactive trajectories (where rare events dictate the dynamics of the system), and the computation of average properties of nonequilibrium systems (which completes the more traditional field of expertise associated with techniques to compute free energy differences).

5.2.1. Sampling of reactive trajectories

Finding trajectories for which the system undergoes a significant change is a challenging task since the transition events are typically very rare. Several methods have been proposed in the physics and chemistry literature, and members of the team have undertaken their study in the past years.

A prominent example is the parallel replica method where several replicas of the system evolve on different processors, until one of them undergoes a transition. Several extensions and refinements to the original method were proposed by T. Lelièvre:

- together with D. Aristoff and G. Simpson, he proposed in [7] an adaptation of the Parallel Replica method for Markov chains;
- together with A. Binder and G. Simpson, he introduced in [17] a generalized parallel replica dynamics. The idea is to extend the applicability of the original algorithm by computing on the fly the so-called decorrelation time.

Another class of techniques to compute reactive trajectories is based on splitting techniques. C.E. Bréhier, T. Lelièvre and M. Rousset have performed in [21] an analysis of the Adaptive Multilevel Splitting algorithm, which is a rare event simulation method where several replicas are evolved concurrently, and selected to favor exploration in a given direction. The computational cost of the algorithm is studied in details in the limit of a large number of replicas.

5.2.2. Nonequilibrium systems

G. Stoltz, together with G. Pavliotis (Imperial College) and Rémi Joubaud, studied in [27] the response of equilibrium systems evolving according to a Langevin dynamics, to external, space-time dependent forcings. In particular, they found out that, even if the external forcing is periodic in time and space with a vanishing space-time average, the systems in general evolves with a non-zero average velocity. It may even be the case that the average velocity is in the direction opposite to the average forcing (when the latter is non-zero), which can be seen as an example of negative mobility. The behavior of the system over diffusive time scales (in the reference frame obtained by removing the average velocity) is also studied, for arbitrary forcing strengths. This work was initiated when G. Pavliotis was a visiting member of the team MATHERIALS.

A numerical analysis of the error arising in the computation of transport coefficients, with an emphasis on mobility and self-diffusion, was provided by M. Fathi, A.A Homman and G. Stoltz in [25] in the case when Metropolis-Hastings algorithms are used to stabilize straightforward discretizations of overdamped Langevin dynamics.

Together with Herbert Spohn (TU München), G. Stoltz has verified the relevance of mode-coupling predictions for the scaling of space time correlations of invariants for one dimensional systems subjected to a non-reversible deterministic dynamics perturbed by an exchange noise [32]. In particular, it has been confirmed that the equilibrium relaxation of the invariants involves two modes, a traveling sound mode and a standing heat mode (related to the energy current and height autocorrelation functions). Both modes exhibit a superdiffusive scaling, of Lévy type for the heat mode, and of KPZ type for the sound mode.

5.2.3. Free energy computations

The topic of free energy computations is still a significant research area of the team. T. Lelièvre and G. Stoltz, together with G. Fort and B. Jourdain, studied the Self-Healing Umbrella Sampling (SHUS) method in [26]. This method is an adaptive biasing method to compute free energies on the fly by appropriately penalizing already visited regions. The convergence of the method relies on a rewriting as a stochastic approximation method with random steps, and can therefore be seen as a variation of the Wang-Landau method. The efficiency of the SHUS algorithm was assessed for a model two-dimensional system in terms of exit times out of a metastable set.

Concerning practical applications, G. Stoltz, together with A.A. Homman, E. Bourasseau, P. Malfreyt, L. Strafella and A. Ghoufi have worked on the computation of surface tension in droplets [10], using alchemical transformations where the droplet volume is artificially varied.

Finally, T. Lelièvre, together with J. Comer, J.C. Gumbart, J. Hénin, A. Pohorille and C. Chipot, wrote a review article on the adaptive biasing force method [9].

5.2.4. Thermodynamic limit

Another work in progress is related to the understanding of the origin of hysteresis in rubber-made materials. When submitted to cyclic deformations, the strain-stress curve of these materials indeed shows a hysteresis behavior, which seems to be independent of the speed of loading.

Some years ago, members of the team have suggested a model, at a mesoscale, to explain this behavior. This model was written in terms of a system made of a finite number of particles. One of the aim of the post-doc of Thomas Hudson, who joined the team in Sept. 2014, is to make progress on that question, and to understand whether a thermodynamic limit of the model previously proposed can be identified.

5.2.5. Reduced models

We propose in [13] a procedure for replacing a complex, reactive potential of REBO type by a simple harmonic approximation, in regions where the system is close to equilibrium. The parameters of the harmonic approximation are chosen so that the phonon spectrum is exactly reproduced. We are currently testing the ability of the so-obtained hybrid model to predict the fracture of graphene.

5.3. Complex fluids

Participants: Sébastien Boyaval, Claude Le Bris, Tony Lelièvre.

Sébastien Boyaval has pursued his research about the mathematical modelling of complex free-surface flows. On the one hand, the numerical investigation of 3D effects with a VOF approach was carried out for multiphase flows in collaboration with the CFSFlow code developers at EPFL [11]. On the other hand, the reduced modelling of viscoelastic effects within Saint-Venant framework was carried out for asymptotically thin layers above rough bottoms [8].

5.4. Application of greedy algorithms

Participants: Sébastien Boyaval, Eric Cancès, Virginie Ehrlacher, Tony Lelièvre.

Model reduction techniques are very important tools for applications. They consist in deriving from a high-dimensional problem, a low-dimensional model, which very quickly gives reliable results. In particular, the team is interested in two techniques: Proper Generalized Decomposition (greedy algorithms) and Reduced Basis techniques.

Eric Cancès, Virginie Ehrlacher and Tony Lelièvre have extended a greedy algorithm suggested for the resolution of high-dimensional eigenvalue problems in order to approximate the solution of the many-body Schrödinger electronic problem. The main technical difficulty in the extension of these algorithms lies in the antisymmetry of the wavefunction of the electrons. To deal with this difficulty, an approximation of the wavefunction is computed as a sum of Slater determinants, each Slater determinant function being computed in an iterative way.

Virginie Ehrlacher has obtained preliminary encouraging results on greedy algorithms for parametric eigenvalue problems. The method has been applied to the computation of the first buckling mode of a plate in the presence of a defect, the position of the defect playing the role of a parameter entering the eigenvalue problem defining the first buckling mode of the plate.

A new numerical method for the construction of an efficient reduced-order model for the solution of the Vlasov equation, arising in plasma physics or in the modeling of electron transport in semiconductors, has been tested by Damiano Lombardi (REO Inria team) and Virginie Ehrlacher. This method is based on the use of an analytic Lax Pair for the Vlasov equations and is inspired by previous works done on transport equations by Jean-Frederic Gerbeau, Damiano Lombardi and Elisa Schenone. Encouraging preliminary numerical results have been obtained.

5.5. Homogenization and related topics

Participants: Sébastien Brisard, Ludovic Chamoin, Virginie Ehrlacher, Claude Le Bris, Frédéric Legoll, Simon Lemaire, François Madiot, William Minvielle.

The homogenization of (deterministic) non periodic systems is a well known topic. Although well explored theoretically by many authors, it has been less investigated from the standpoint of numerical approaches (except in the random setting). In collaboration with X. Blanc and P.-L. Lions, C. Le Bris has introduced a possible theory, giving rise to a numerical approach, for the simulation of multiscale nonperiodic systems. The theoretical considerations are based on earlier works by the same authors (derivation of an algebra of functions appropriate to formalize a theory of homogenization). The numerical endeavour is completely new. The theoretical results obtained to date are being collected in a series of manuscripts that will be available shortly.

The team has pursued its efforts in the field of stochastic homogenization of elliptic equations, aiming at designing numerical approaches that both are pratically relevant and keep the computational workload limited.

Using the standard homogenization theory, one knows that the homogenized tensor, which is a deterministic matrix, depends on the solution of a stochastic equation, the so-called corrector problem, which is posed on the *whole* space $\mathbb{R}^d$. This equation is therefore delicate and expensive to solve. In practice, the space $\mathbb{R}^d$ is truncated to some bounded domain, on which the corrector problem is numerically solved. In turn, this yields a converging approximation of the homogenized tensor, which happens to be a *random* matrix.

In [28], F. Legoll and W. Minvielle have proposed a variance reduction procedure, based on the control variate technique, to obtain estimates of the apparent homogenized tensor with a smaller statistical error (at a given computational cost) than standard Monte Carlo approaches. The control variate technique is based on using a surrogate model, somewhat in the spirit of a preconditionner. In [28], the surrogate model that is used is inspired by a weakly stochastic approach previously introduced by A. Anantharaman and C. Le Bris to describe periodic models perturbed by rare defects.

In addition, C. Le Bris, F. Legoll and W. Minvielle have investigated the possibility to use another variance reduction technique based on computing the corrector equation only for selected environments. These environments are chosen based on the fact that their statistics in the finite supercell matches the statistics of the materials in the infinite supercell. This method yields an estimator with a smaller variance that standard estimators. Preliminary encouraging numerical results have been obtained.

As pointed out above, the corrector problem is in practice solved on a large bounded domain, often complemented with periodic boundary conditions. Solving that problem can still be challenging, in particular because producing a conforming mesh of realistic heterogeneous microstructures can be a daunting task. In such situations, numerical methods formulated on cartesian grids may be more interesting. These methods can still be Finite Element Methods, or methods in the spirit of that proposed by Moulinec and Suquet in the mid-nineties. In their approach, the corrector problem (a partial differential equation) is reformulated as an equivalent integral equation. This equation can readily be discretized using a Galerkin approach. This leads to numerical schemes that can be implemented as a matrix-free method. In [18], S. Brisard and F. Legoll have reviewed the different variants that have been proposed in the literature along these ideas, and proposed a mathematical analysis of the numerical schemes. This work extends in various directions previous works by S. Brisard.

In somewhat the same vein, Eric Cancès, Virginie Ehrlacher and Frédéric Legoll (in collaboration with Benjamin Stamm, University Paris 6) have worked on alternative methods to approximate the homogenized coefficients of a random stationary material. These methods are alternative to those proposed e.g. by Bourgeat and Piatniski, and which consist in solving a corrector problem on a bounded domain. The method introduced is based on a new corrector problem. This problem is posed on the entire space. In some cases (including the case of randomly located spherical inclusions), it can be recast as an integral equation posed on the surface of the inclusions. The problem can then be efficiently solved via domain decomposition and using spherical harmonics.

We have discussed above approaches to efficiently compute the homogenized coefficient, assuming we have a complete knowledge of the microstructure of the material. We have actually also considered a related inverse problem, and more precisely a parameter fitting problem. Knowing the homogenized quantities, is it possible to recover some features of the microstructure properties? Obviously, since homogenization is an averaging procedure, not everything can be recovered from macroscopic quantities. A realistic situation is the case when a functional form of the distribution of the microscopic properties is assumed, but with some unknown parameters to determine. In collaboration with A. Obliger and M. Simon, F. Legoll and W. Minvielle have addressed that problem in [29], showing how to determine the unknown parameters of the microscopic distribution on the basis of macroscopic (e.g. homogenized) quantities.

From a numerical perspective, the Multiscale Finite Element Method (MsFEM) is a classical strategy to address the situation when the homogenized problem is not known (e.g. in difficult nonlinear cases), or when the scale of the heterogeneities, although small, is not considered to be zero (and hence the homogenized problem cannot be considered as an accurate enough approximation).

The MsFEM has been introduced more than 10 years ago. However, even in simple deterministic cases, there is actually still room for improvement in many different directions. In collaboration with A. Lozinski (University of Besançon), F. Legoll and C. Le Bris have introduced and studied a variant of MsFEM that considers Crouzeix-Raviart type elements on each mesh element. The continuity across edges (or facets) of the (multiscale) finite element basis set functions is enforced only weakly, using fluxes rather than point values. That approach has been analyzed and tested on an elliptic problem set on a domain with a huge number of perforations. The variant developed outperforms all existing variants of MsFEM.

A follow up on this work, in collaboration with U. Hetmaniuk (University of Washington in Seattle) and A. Lozinski (University of Besançon), consists in the study of multiscale advection-diffusion problems. Such problems are possibly advection dominated and a stabilization procedure is therefore required. How stabilization interplays with the multiscale character of the equation is an unsolved mathematical question worth considering for numerical purposes. In that spirit, C. Le Bris, F. Legoll and F. Madiot have studied several variants of the Multiscale Finite Element Method (MsFEM), specifically designed to address multiscale advection-diffusion problems in the convection-dominated regime. Generally speaking, the idea of the MsFEM is to perform a Galerkin approximation of the problem using specific basis functions, that are precomputed (in an offline stage) and adapted to the problem considered. Several possibilities for the basis functions have been examined (for instance, they may or may not encode the convection field). The various approaches have been compared in terms of accuracy and computational costs.

Most of the numerical analysis studies of the MsFEM are focused on obtaining *a priori* error bounds. In collaboration with L. Chamoin, who is currently in delegation in our team (from ENS Cachan, since September 2014), we have started to work on *a posteriori* error analysis for MsFEM approaches, with the aim to develop error estimation and adaptation tools. We have extended to the MsFEM case an approach that is classical in the computational mechanics community for single scale problems, and which is based on the so-called Constitutive Relation Error (CRE). Once a numerical solution u_h has been obtained, the approach needs additional computations in order to determine a divergence-free field as close as possible to the exact flux $k\nabla u$. In the context of the MsFEM, it is important to be able to do all the expensive computations in an offline stage, independently of the right-hand side. The standard CRE approach thus needs to be adapted to that context, in order to keep that feature that makes it adapted to a multiscale, multi-query context. The preliminary approach that we have introduced already yields promising results.

Still another question investigated in the group is to find an alternative to standard homogenization techniques when these latter are difficult to use in practice. This is the aim of the post-doc of Simon Lemaire, which began in June 2014, and which takes over previous works of the group on the subject. Consider a linear elliptic equation, say in divergence form, with a highly oscillatory matrix coefficient, and assume that this problem is to be solved for a large number of right-hand sides. If the coefficient oscillations are infinitely rapid, the solution can be accurately approximated by the solution to the homogenized problem, where the homogenized coefficient has been evaluated beforehand by solving the corrector problem. If the oscillations are moderately rapid, one can think instead of MsFEM-type approaches to approximate the solution to the reference problem. However, in both cases, the complete knowledge of the oscillatory matrix coefficient is required, either to build the average model or to compute the multiscale basis. In many practical cases, this coefficient is often only partially known, or merely completely unavailable, and one only has access to the solution of the equation for some loadings. This observation has lead to think about alternative methods, in the following spirit. Is it possible to approximate the reference solution by the solution to a problem with a *constant* matrix coefficient? How can this 'best' constant matrix approximating the oscillatory problem be constructed in an efficient manner?

A preliminary step, following discussion and interaction with A. Cohen, has been to cast the problem as a convex optimization problem. We have then shown that the 'best' constant matrix defined as the solution of that problem converges to the homogenized matrix in the limit of infinitely rapidly oscillatory coefficients. Furthermore, the optimization problem being convex, it can be efficiently solved using standard algorithms. C. Le Bris, F. Legoll and S. Lemaire are currently working on making the resolution of the optimization problem as efficient as possible.

To conclude this section, we mention a project involving V. Ehrlacher, C. Le Bris and F. Legoll, in collaboration with G. Leugering and M. Stingl (Cluster of Excellence, Erlangen-Nuremberg University). This project aims at optimizing the shape of some materials (modelled as structurally graded linear elastic materials) in order to achieve the best mechanical response at the minimal cost. As often the case in shape optimization, the solution tends to be highly oscillatory, thus the need of homogenization techniques. Materials under consideration are being thought of as microstructured materials composed of steel and void and whose microstructure patterns are constructed as the macroscopic deformation of a reference periodic microstructure. The optimal material (i.e. the best macroscopic deformation) is the deformation achieving the best mechanical response. For a given deformation, we have first chosen to compute the mechanical response using a homogenized model. We are currently aiming at computing the mechanical response at the microscale, using the highly oscillatory model. Model reduction techniques (such as MsFEM, Reduced Basis methods, ...) are then in order, in order to expedite the resolution of the oscillatory problem, which has to be solved at each loop of the optimization algorithm. Current efforts are targeted towards choosing an appropriate model reduction strategy.

5.6. Miscellaneous

Participants: Sébastien Boyaval, Tony Lelièvre, Sébastien Boyaval.

T. Lelièvre together with F. Casenave and A. Ern propose in [24] an extension of the classical reduced basis method in order to extend its range of applicability to black-box codes.

S. Boyaval started investigating new high-order methods on generalized non-conforming meshes in collaboration with Daniele di Pietro [14].

In [31], M. Rousset considers space homogenous Boltzmann kinetic equations in dimension $d \geq 3$ with Maxwell collisions (and without Grad's cut-off). An explicit Markov coupling of the associated conservative stochastic N-particle system is constructed, yielding a N-uniform α-power law trend to equilibrium.

MATHRISK Project-Team

6. New Results

6.1. Highlights of the Year

B. Jourdain and A. Sulem : Guest editors of the special issue "Systemic Risk" of *Statistics and Risk Modeling*, 2014. [27]

The research project "Stochastic Control of Systemic Risk" has been awarded by the scientific council and Professional Fellows of Institut Europlace de Finance (EIF) and Labex Louis Bachelier (December 2014).

Roxana Dumitrescu, PhD student, received the price for collaborative actions during her PhD studies, delivered by Fondation des Sciences Mathématiques de Paris and CASDEN (November 2014).

Pierre Blanc, PhD student, has got the award of "Rising star of quantitative finance" for his talk on a price impact models with an exogeneous (Hawkes) flow of orders [29]. This prize was given by the Global Derivatives conference (Amsterdam, 12-16 May) to indicate the best work among PhD students.

6.2. Liquidity risk

Aurélien Alfonsi and his PhD student Pierre Blanc are working on the optimal execution problem when there are many large traders who modify the price. They consider an Obizhaeva and Wang model for the price impact, and they assume that the flow of market orders generated by the other traders is given by an exogenous process. They have shown that Price Manipulation Strategies (PMS) exist when the flow of order is a compound Poisson process. On the other hand, modeling this flow by a mutually exciting Hawkes process allows them with a particular parametrization to exclude these PMS. Besides, they are able to calculate explicitly the optimal execution strategy within the model [29]. They are now investigating how this model can fit market data.

6.3. Dependence modeling

With his PHD student J. Reygner, B. Jourdain has studied a mean-field version of rank-based models of equity markets, introduced by Fernholz in the framework of stochastic portfolio theory ([38]). When the number of companies grows to infinity, they obtain an asymptotic description of the market in terms of a stochastic differential equation nonlinear in the sense of McKean. The diffusion and drift coefficients depend on the cumulative distribution function of the current marginal law of the capitalizations. Using results on the longtime behavior of such SDEs derived in [66], they discuss the long-term capital distribution in this asymptotic model, as well as the performance of simple portfolio rules. In particular, they highlight the influence of the volatility structure of the model on the growth rates of portfolios.

Another approach to handle the question of stochastic modeling in a multidimensional framework consists in dealing with stochastic differential equations that are defined on matrices in order to model either the instantaneous covariance or the instantaneous correlation between the assets.

The research on the estimation of the parameters of a Wishart process has started this year together with the thesis of Clément Rey. A. Alfonsi, A. Kebaier and C. Rey are studying the Maximum Likelihood Estimator for the Wishart processes and in particular its convergence in the ergodic and the non ergodic case.

Correlation issues are crucial in the modeling of volatility. In his thesis, Ould Aly ([77]) proposes a revised version of Bergomi's model for the variance curve which proves to be very tractable for calibration and for the pricing of variance derivatives (see [23]). He also obtains results on the monotonicity of option prices with respect to the correlation between the stock price and the volatility in the Heston model (see [78]).

In [34], [15], L. Abbas-Turki and D. Lamberton study the monotonicity of option prices with respect to cross-asset correlations in a multidimensional Heston model.

Modeling the dependence is not only useful for the equity market. In credit risk, getting a model that describes the dynamic of the joint distribution of a basket of defaults is still a challenge. The Loss Intensity model proposed by Schönbucher allows to fit perfectly the marginal distributions of the number of defaults in a basket. Then, Stochastic Loss Intensity models extend this model and can also in principle fit the marginal distributions. However, these models appear as a non-linear differential equation with jumps. A Alfonsi, C. Labart and J. Lelong have shown that these models are well-defined by using a particles system ([44]). Besides, this particles system gives a very convenient way to run a Monte-Carlo algorithm and to compute expectations in this model. Interacting particle systems are studied by B. Jourdain and his PhD student Julien Reygner in [39], [21].

Application of optimal transport. A. Alfonsi and B. Jourdain study in [43] the Wasserstein distance between two probability measures in dimension n sharing the same copula C. The image of the probability measure dC by the vectors of pseudo-inverses of marginal distributions is a natural generalization of the coupling known to be optimal in dimension $n = 1$. In dimension $n > 1$, it turns out that for cost functions equal to the p-th power of the L^q norm, this coupling is optimal only when $p = q$ i.e. when the cost function may be decomposed as the sum of coordinate-wise costs.

6.4. Systemic risk

The mathematical modeling of default contagion, by which an economic shock causing initial losses and default of a few institutions is amplified due to complex linkages, leading to large scale defaults, can be addressed by various techniques, such as network approaches (see in particular [46]), or mean field interaction models [62], [55]. Little has been done so far on the *control* of such systems and A. Sulem has started to contribute on these issues in the framework of random graph models in collaboration with A. Minca (Cornell University) and H. Amini (EPFL). In [22], [31], they consider a financial network described as a weighted directed graph, in which nodes represent financial institutions and edges the exposures between them. Here, the distress propagation is modeled as an epidemics on this graph. They study the optimal intervention of a lender of last resort who seeks to make equity infusions in a banking system prone to insolvency and to bank runs, under complete and incomplete information of the failure cluster, in order to minimize the contagion effects.

R. Elie is studying risk systemic propagation and its links with mean field games.

6.5. Backward stochastic (partial) differential equations with jumps and stochastic control with nonlinear expectation

A. Sulem, M.C. Quenez and R. Dumitrescu have studied optimization problems for BSDEs with jumps [11], optimal stopping for dynamic risk measures induced by BSDEs with jumps and associated reflected BSDEs. [24], [80], [19]. They have also investigated optimal stopping with nonlinear expectation under ambiguity, and their links with nonlinear Hamilton Jacobi Bellman variational inequalities in the Markovian case. Moreover they have obtained dynamic programming principles for mixed optimal-stopping problems with nonlinear expectations. They have also explored the links between generalized Dynkin games and double barriers reflected BSDE with jumps [56]. Stochastic control of Itô-Lévy Processes with applications to finance are studied by A. Sulem and B. Øksendal in [25], [26]. We have also contributed to the theory of BSDEs and Forward-Backward SDEs which appear as the adjoint equations associated to stochastic maximum principles, and address various issues about the relation between information and performance in non Markovian stochastic control: In particular, in the context of jump-diffusion models under partial information, A. Sulem, C. Fontana and B. Øksendal study in [20] the relation between market viability (in the sense of solvability of portfolio optimization problems) and the existence of a martingale measure given by the marginal utility of terminal wealth, without a-priori assuming no-arbitrage restrictions on the model.

A. Sulem, with B. Øksendal and T. Zhang has studied optimal stopping for Stochastic Partial Differential equations and associated reflected SPDEs [91], and optimal control of Forward-Backward SDEs [90].

Stochastic maximum principles for singular mean-field games are obtained in [37] with applications to optimal irreversible investments under uncertainty.

R. Dumitrescu and C. Labart have proposed a numerical approximation for Doubly Reflected BSDEs with Jumps and RCLL obstacles [35].

R. Elie studies approximate hedging prices under various risk constraints. This is done in collaboration with P. Briand, Y. Hu, A. Matoussi, B. Bouchard, L. Moreau, J.F. Chassagneux, I. Kharroubi and R. Dumitrescu.

6.6. Option Pricing

Interest rates modeling. A. Alfonsi studies an affine term structure model for interest rates that involve Wishart diffusions (with E. Palidda) [28] . Affine term structure models (Dai and Singleton, Duffie, ...) consider vector affine diffusions. Here, we extend the Linear Gaussian Model (LGM) by including some Wishart dynamics, and to get a model that could better fit the market. We have obtained a price expansion around the LGM for Caplet and Swaption prices. Also, we present a second order discretization scheme that allow to calculate exotic prices with this model.

American Options. In joint work with Aych Bouselmi, D. Lamberton studied the asymptotic behavior of the exercise boundary near maturity for American put options in exponential Lévy models [34].

He is currently working with M. Pistorius on the approximation of American options by Canadian options, which originated from the work of Peter Carr.

Barrier Options. Numerical pricing of double barrier options is investigated by A. Zanette and coauthors in [16].

6.7. Discretization of stochastic differential equations

With his PhD student A. Al Gerbi and E. Clément, B. Jourdain is interested in the strong convergence properties of the Ninomiya-Victoir scheme which is known to exhibit order 2 of weak convergence. This study is aimed at analysing the use of this scheme either at each level or only at the finest level of a multilevel Monte Carlo estimator : indeed, the variance of a multilevel Monte Carlo estimator is related to the strong error between the two schemes used in the coarse and fine grids at each level. They prove strong convergence with order $1/2$ which is improved to order 1 when the vector fields corresponding to each Brownian coordinate in the SDE commute. They also check that the renormalized errors converge to affine SDEs with source terms involving the Lie brackets between these vector fields and, in the commuting case, their Lie brackets with the drift vector field. Last, they propose a modified Ninomiya-Victoir scheme, which, at the finest level of the multilevel Monte Carlo estimator, may be coupled with strong order 1 to a simpler scheme with weak order 1 recently proposed by Giles and Szpruch.

Using optimal transport tools, A. Alfonsi, B. Jourdain and A. Kohatsu-Higa have proved that the Wasserstein distance between the time marginals of an elliptic SDE and its Euler discretization with N steps is not larger than $\frac{C\sqrt{\log(N)}}{N}$. The logarithmic factor may be removed when the uniform time-grid is replaced by a grid still counting N points but refined near the origin of times [4] . To generalize in higher dimension the result that they obtained previously in dimension one using the optimality of the explicit inverse transform, they compute the derivative of the Wasserstein distance with respect to the time variable using the theory developed by Ambrosio Gigli and Savare. The abstract properties of the optimal coupling between the time marginals then enable them to estimate this time derivative [30].

6.8. Advanced Monte Carlo methods.

- **Adaptive variance reduction methods.** B. Jourdain and J. Lelong have pursued their work on adaptive Monte Carlo methods in several directions [17], [36].

- **Metropolis Hastings algorithm in large dimension.** With T. Lelièvre and B. Miasojedow, B. Jourdain considers the Random Walk Metropolis algorithm on $\mathbf{R}^n$ with Gaussian proposals, and when the target probability measure is the n-fold product of a one dimensional law. It is well-known that, in the limit n tends to infinity, starting at equilibrium and for an appropriate scaling of the variance and of the timescale as a function of the dimension n, a diffusive limit is obtained for each component of the Markov chain. They generalize this result when the initial distribution is not the target probability measure ([65]). The obtained diffusive limit is the solution to a stochastic differential equation nonlinear in the sense of McKean. In [64], they prove convergence to equilibrium for this equation. They also discuss practical counterparts in order to optimize the variance of the proposal distribution to accelerate convergence to equilibrium. The analysis confirms the interest of the constant acceptance rate strategy (with acceptance rate between 1/4 and 1/3).

6.9. Numerical Probability

6.9.1. Regularity of probability laws using an interpolation method

This work was motivated by previous studies by N. Fournier, J. Printemps, E. Clément, A. Debusche and V. Bally, on the regularity of the law of the solutions of stochastic differential equations with low regularity coefficients - such as diffusion processes with Hölder coefficients or many other examples including jump type equations, Boltzmann equation or Stochastic PDE's. Since we do not have sufficient regularity, the usual approach by Malliavin calculus fails in this framework. We use the following alternative idea: We approximate the law of the random variable X (the solution of the equation at hand) by a sequence $X(n)$ of random variables which are smooth. Consequently we are able to establish integration by parts formulas for $X(n)$, to obtain the absolutely continuity of the law of $X(n)$, and to establish estimates for the density of the law of $X(n)$ and its derivatives. Note that the derivatives of the densities of $X(n)$ generally blow up - so we can not derive directly results concerning the density of the law of X. But, if the speed of convergence of $X(n)$ to X is faster than the blow up, then we may obtain results concerning the density of the law of X. It turns out that this approach fits in the framework of interpolation spaces and that the criterion of regularity for the law of X amounts to the characterization of an interpolation space between a space of distributions and a space of smooth functions. Although the theory of interpolation spaces is very well developed and one already knows how to characterize the interpolation spaces for Sobolev spaces of positive and negative indices, we have not found in the (huge) literature a result which covers the problem we are concerned with. So, although our result may be viewed as an interpolation result, it is a new one. The above work is treated in the paper [48] by V. Bally and Lucia Caramellino. As an application we discussed in [50] the regularity of the law of a Wiener functional under a Hörmander type non degeneracy condition.

6.9.2. A stochastic parametrix representation for the density of a Markov process.

Classical results of PDE theory (due to A. Friedmann) assert that, under uniform ellipticity conditions, the law of a diffusion process has a continuous density (the approach of A. Friedmann is analytical and concerns PDE's instead of the corresponding diffusion process). The method developed by A. Friedmann is known as the "parametrix method". V. Bally In collaboration with A. Kohatzu Higa gave a probabilistic approach which represents the probabilistic counterpart of the parametrix method [33]. They obtained a probabilistic representation for the density of the law of the solution of a SDE and more generally, for a class of Markov processes including solutions of jump type SDE's. This representation may be considered as a perfect simulation scheme and so represents a starting point for Monte Carlo simulation. However the random variable which appears in the stochastic representation has infinite variance, so direct simulation gives unstable results (as some preliminary tests have proved). In order to obtain an efficient simulation scheme some more work on the reduction of variance has to be done - and this does not seem trivial.

6.9.3. The distance between two density functions and convergence in total variation.

V. Bally and L. Caramellino have obtained estimates of the distance between the densities of the law of two random variables using an abstract variant of Malliavin calculus. They used these estimates in order to study the convergence in total variation of a sequence of random variables. This has been done in [49]. They are now working on more specific examples concerning the Central Limit Theorem [32]. In the last years the convergence in entropy distance and in total variation distance for several variants of the CLT has been considered in papers by S. Bobkov, F. Götze, G. Peccati, Y. Nourdin, D. Nualart and G. Poly. This is a very active research. Moreover, in an working paper in collaboration with his Phd student R. Clement, V. Bally uses similar methods in order to study the total variation distance between two Markov semigroups and for approximation schemes purposes. A special interest is devoted to higher order schemes such as the Victoir Nyomia scheme.

6.9.4. An invariance principle for stochastic series (U- Statistics).

Vlad Bally and Lucia Caramellino are working on invariance principles for stochastic series of polynomial type. In the case of polynomials of degree one we must have the classical Central Limit Theorem (for random variables which are not identically distributed). For polynomials of higher order we are in the framework of the so called U statistics which have been introduced by Hoffdings in t 1948 and which play an important role in modern statistics. Our contribution in this topic concerns convergence in total variation distance for this type of objects. We use abstract Malliavin calculus and more generally, the methods mentioned in the above paragraph.

MAVERICK Project-Team

5. New Results

5.1. Highlights of the Year

The impacting PhD work [3] of Eric Heitz on *appearance filtering* (see section 5.5.1) has received a very good reception in both academic and industrial world, including several “best paper” prizes in 2013 and 2014, invitation to participate to the Siggraph Course on Photorealistic Rendering [13], and statements of importance and/or integration by reference peoples and CG companies.

BEST PAPER AWARD :

[] **Computer Graphics Forum**. E. HEITZ, E. D'EON.

5.2. Visual perception

5.2.1. *The effects of surface gloss and roughness on color constancy for real 3-D objects*

Participants: Jeoren J. M. Granzier, Romain Vergne [contact], Karl Gegenfurtner.

Color constancy denotes the phenomenon that the appearance of an object remains fairly stable under changes in illumination and background color. Most of what we know about color constancy comes from experiments using flat, matte surfaces placed on a single plane under diffuse illumination simulated on a computer monitor. Here we investigate whether material properties (glossiness and roughness) have an effect on color constancy for real objects. Subjects matched the color and brightness of cylinders (painted red, green, or blue) illuminated by simulated daylight (D65) or by a reddish light with a Munsell color book illuminated by a tungsten lamp. The cylinders were either glossy or matte and either smooth or rough. The object was placed in front of a black background or a colored checkerboard as shown in Figure 6 . We found that color constancy was significantly higher for the glossy objects compared to the matte objects, and higher for the smooth objects compared to the rough objects. This was independent of the background. We conclude that material properties like glossiness and roughness can have significant effects on color constancy [7].

5.3. Visualization

Participants: Léo Allemand-Giorgis, Georges-Pierre Bonneau [contact].

In computer visualization we have worked on two topics: topology for visualization and perception for visualization.

In topology for visualization we have worked on scalar field vizualization methods taking into account the topology of the data. In [14] We have derived theoretical results on monotonic interpolation of scalar data. Our method enables to interpolate given topological data such as minima, maxima and saddle points at the corners of a rectangular domain without adding spurious extrema inside the function domain, as illustrated in Figure 7 .

We have collaborated to a state of the art chapter on Uncertain Visualization [15], in which we described the evaluation of visualization methods based on visual perception.

Furthermore we have worked on two topics related to geometry for visualization. In [6] we introduce a method for interpolating a quad mesh using G1-continuous polynomial surfaces. We plan to use this method in the future for displaying isosurfaces of higher order data. In [11] we have published a method for reconstructing interfaces in highly complex assemblies, as illustrated in Figure 8 . This method has been developed in order to visualize data arising from simulation of complex mechanical assemblies, within the ANR project ROMMA, closed in January 2014.

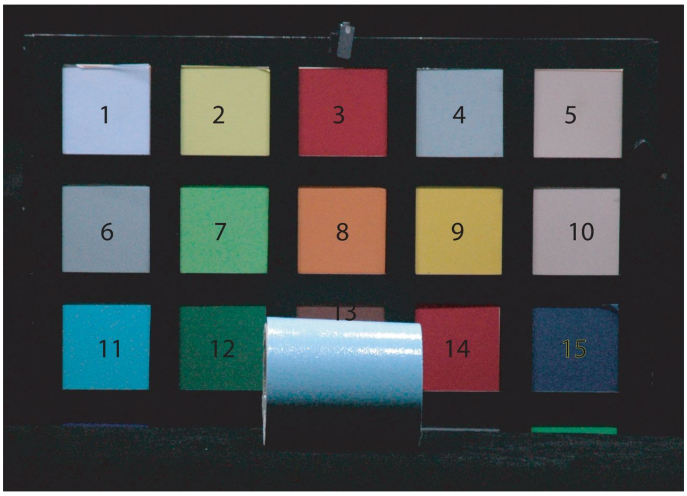


Figure 6. Color perception depends on material properties. This image represents one stimulus used in our experiment to compare the effect of glossiness on color constancy.

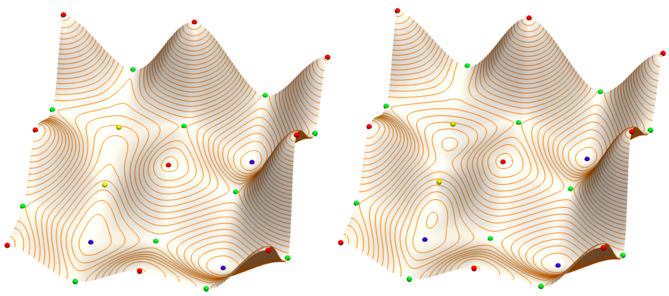


Figure 7. Local maxima (red), minima (blue), saddles (green) and regular (yellow) vertices are interpolated by a C1 piecewise cubic interpolant. Left: no unwanted local extrema exist in the interior of the cubic patches. Right: partial derivatives too large in size are chosen for the yelllow regular vertices implying that additional unwanted local extrema appear inside the cubic polynomial patches.

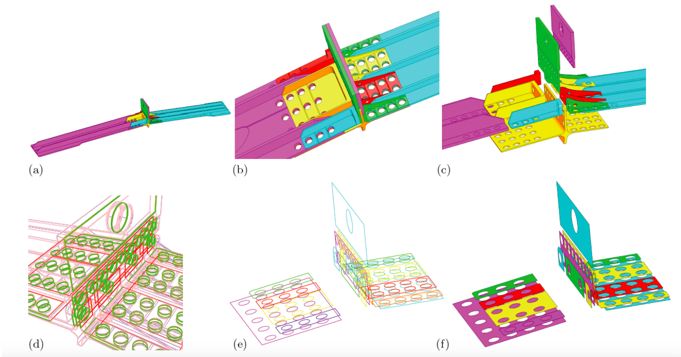


Figure 8. Aircraft part for assembling the wings with the body of an aircraft (model courtesy of EADS). (a,b) two views of the components, (c) exploded view, (d) ray casting results, (e) boundary reconstruction, (f) nal interfaces

5.4. Image creation and editing

5.4.1. Programmable 2D Arrangements for Element Texture Design

Participants: Hugo Loi, Thomas Hurtut, Romain Vergne, Joëlle Thollot [contact].

We introduce a programmable method for designing stationary 2D arrangements for element textures, namely textures made of small geometric elements. These textures are ubiquitous in numerous applications of computer-aided illustration. Previous methods, whether they be example-based or layout-based, lack control and can produce a limited range of possible arrangements. Our approach targets technical artists who will design an arrangement by writing a script. These scripts are using three types of operators: *partitioning operators* for defining the broad-scale organization of the arrangement, *mapping operators* for controlling the local organization of elements, and *merging operators* for mixing different arrangements. These operators are designed so as to guarantee a stationary result meaning that the produced arrangements will always be repetitive. We show (see Figure 10) that this simple set of operators is sufficient to reach a much broader variety of arrangements than previous methods. Editing the script leads to predictable changes in the synthesized arrangement, which allows an easy iterative design of complex structures. Finally, our operator set is extensible and can be adapted to application-dependent needs.

5.4.2. Color transfer guided by summary statistics

Participants: Benoît Arbelot, Romain Vergne [contact], Thomas Hurtut, Joëlle Thollot.

Modifying the colors of an image is an attractive way to edit its ambiance and mood. In practice, manually and directly tuning the color distribution of an image is challenging and tedious. Color transfer methods offer an intuitive alternative by automatically changing an image colors according to a target image. Existing transfer methods mostly rely on global matching processes to reshape and map the color histogram of the source image as close as possible to the target histogram. However, they offer no control over where the colors of the target will be transferred in the source image: they only tend to match colors that have similar intensities and chromaticities. This can lead to unexpected results, especially when some elements do not have the same colors in the two images, but share similar features. In this work, we propose to implicitly segment input images before transferring colors. Instead of relying on colors only, we use a summary of statistics to describe the underlying texture properties of each pixel. This provides a measure of pixel similarity which is then used to guide and ensure the transfer to be done between similar features (see Figure reffig:color for a preliminary result).

Source Target Tranfer

Figure 9. By transferring colors between statistically similar pixels, meaningful colors are transferred from the target to the source image.

5.5. Complex scenes

In order to render both efficiently and accurately ultra-detailed large scenes, this approach consists in developing representations and algorithms able to account compactly for the quantitative visual appearance of a regions of space projecting on screen at the size of a pixel.

5.5.1. Surfacic appearance pre-filtering

Participants: Eric Heitz, Fabrice Neyret [contact].

Here, we deal with complex surfaces represented by microfacets and material attributes.

Among the various correlations between material ingredients forming the BRDF, we published an extended version of the work on correlation between surface attribute (like color) and visibility [9], and Eric published an comprehensive interpretation of the microfacet model in a journal in the field of physics [8]. He also adapted his microfacet approach to the efficient BRDF sampling for path tracing – published in EGSR/CGF [] –, see Figure 11 , and he was invited to participate to the prestigious Siggraph Course “Physically Based Shading in Theory and Practice” [13]. This work is now implemented in various professional and standard software and thus settled a new standard. Eric defended his PhD on September, 26 2014 [3]

5.5.2. Volumetric appearance pre-filtering

Participants: Guillaume Loubet, Fabrice Neyret [contact].

Here, we deal with complex density distributions. The first target is galactic material in the scope of the veRTIGE / Galaxy ANR project, but the long term goal is more general since at long distance complex surfaces or scattered objects can more efficiently be represented as volumetric distributions.

The usual hypothesis in CG is that volumes are homogeneous distribution of matter. But star and (dark) dust distributions are fractal, not homogeneous. This breaks all the existing equations accounting for large scale opacity and lighting of volumes of such material.

first, we developed a new procedural noise able to easily mimic such fractal distributions according to astrophysical models (see fig 12 ,a). Then we studied how to reproduce the same opacity (fig 12 ,b) and reflectivity (fig 12 ,c) for various level of details – this is still ongoing work.

Moreover, volumetric material is often concentrated into bodies, with a boundary delimited by a density jump or gradient. We studied the macroscopic light behavior in such configurations (fig 12 ,d).

5.6. Realistic rendering

Note that Cyril Soler defended his HDR “*Models and Analyses for Image Synthesis*”, Université Joseph-Fourier, on June 2014.

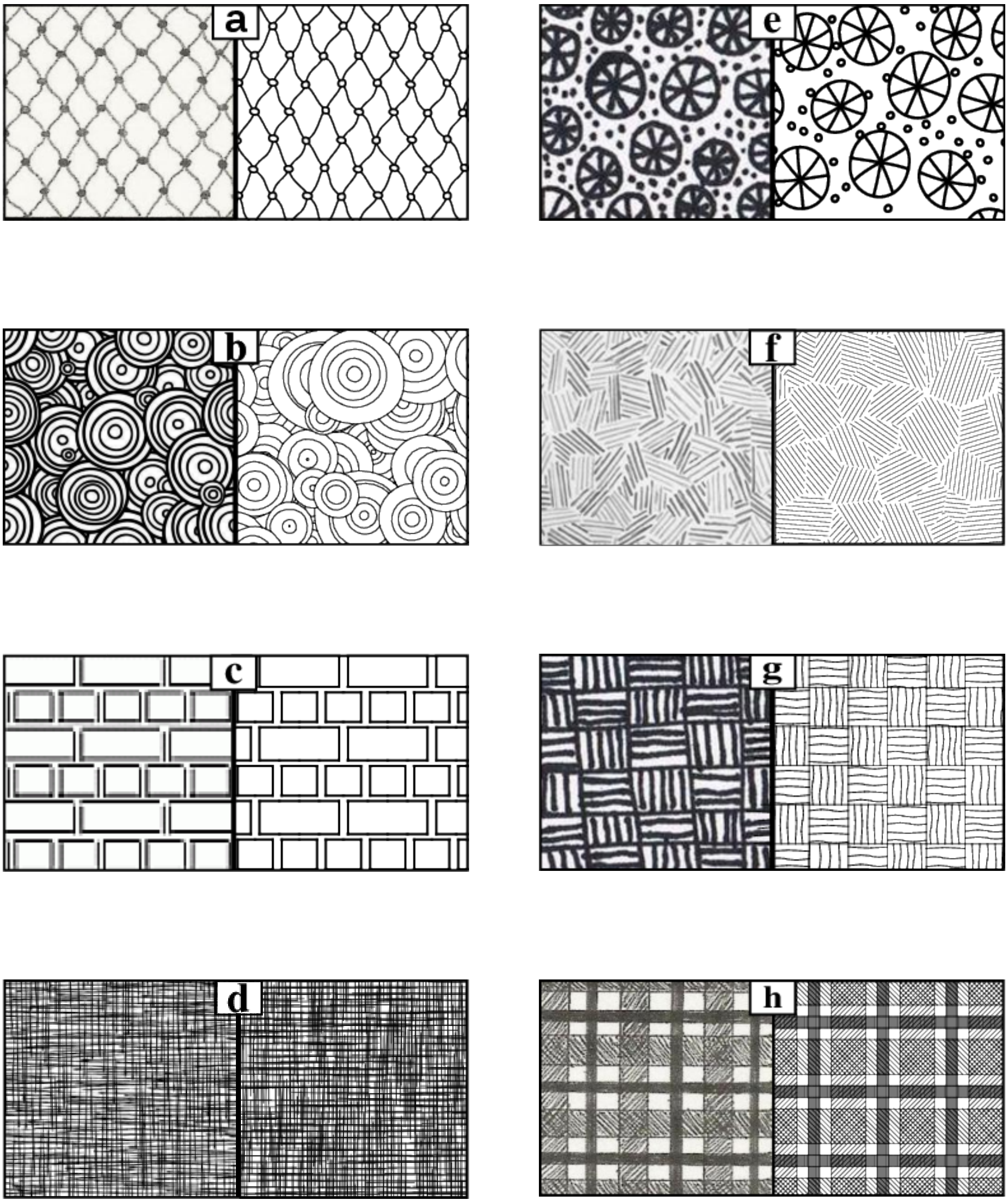


*Figure 10. **Element textures commonly used.** These textures can be found in professional art (d,g,h), casual art (a,e,f), technical productions such as Computer-Assisted Design illustration tools (c), and textile industry (b). For each example, we show a hand-drawn image (left), and our synthesized reproduction of its geometric arrangement (right). **(a,b,c)** Classic regular distributions with contact, overlap and no adjacency between elements respectively. **(d)** Overlap of two textures creating cross hatching. **(e)** Non overlapping combination of two textures. **(f,g,h)** Complex element textures with clusters of elements. — Image credit: (d,g,h) "Rendering in Pen and Ink: The Classic Book On Pen and Ink Techniques for Artists, Illustrators, Architects, and Designers" [20]; (a,e) Profusion Art [profusionart.blogspot.com]; (f) Hayes' Art Classes [hayesartclasses.blogspot.com]; (c) CompugraphX [www.compugraphx.com]; (b) 123Stitch [www.123stitch.com].*

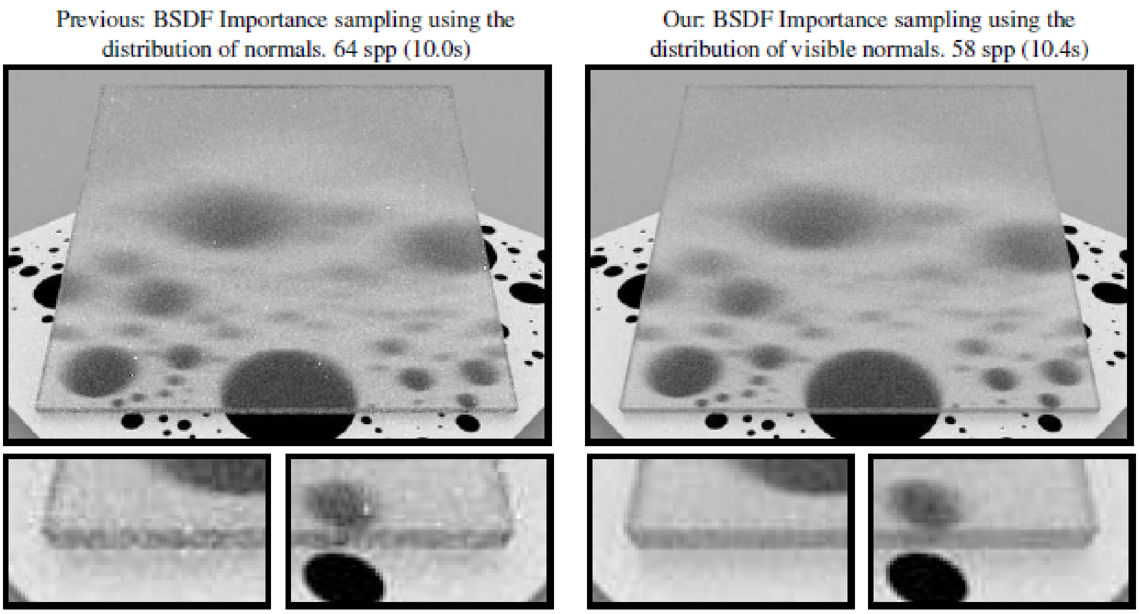


Figure 11. A dielectric glass plate (n = 1.5) with anisotropic GGX roughness (ax = 0.05, ay = 0.4) on all faces (with the Smith masking function). For a similar sample budget and the same render time, our method (right) significantly reduces the variance and converges faster than the common technique used in previous work (left).

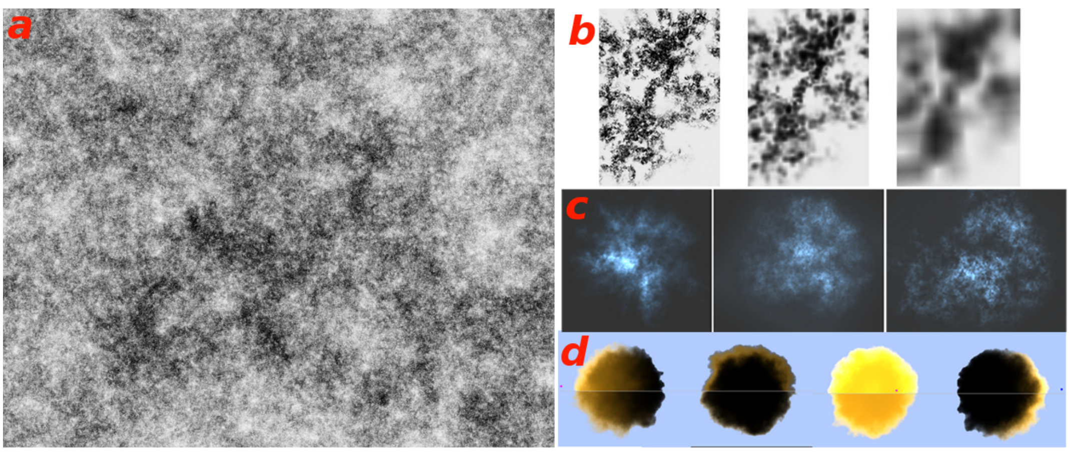
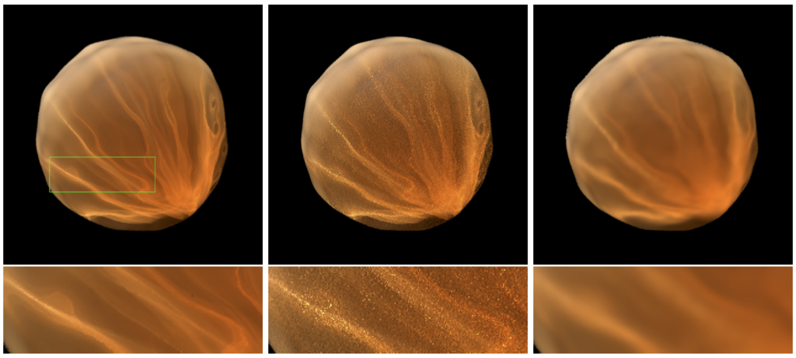


Figure 12. a: Our new fractal procedural noise. b: Multiscale opacity. c: Multiscale reflectance. d: Light reflection at volumetric bodies boundary with gradient (top) or jump (bottom) density distribution, with different light direction (left to right).

5.6.1. Single Scattering in participating media with refractive boundaries

Participant: Nicolas Holzschuch [contact].



(a) Our algorithm (169 s) **(b)** Point samples [WZHB09] (27 samples, 170 s) **(c)** Photon mapping (650 000 photons, 167 s)

Figure 13. Single scattering: comparison between our algorithm and existing methods (equal computation time) on a translucent sphere illuminated by a point light source from behind.

Volume caustics are high-frequency effects appearing in participating media with low opacity, when refractive interfaces are focusing the light rays (see Figure 13). Refractions make them hard to compute, since screen locality does not correlate with spatial locality in the medium. We have developed a new method for accurate computation of single scattering effects in a participating media enclosed by refractive interfaces. Our algorithm is based on the observation that although radiance along each camera ray is irregular, contributions from individual triangles are smooth. Our method gives more accurate results than existing methods, faster. It uses minimal information and requires no precomputation or additional data structures. This paper was accepted for publication at Computer Graphics Forum [10].

5.6.2. A Local Frequency Analysis of Light Scattering and Absorption

Participants: Laurent Belcour, Kavita Bala, Cyril Soler [contact].

We proposed a novel analysis of absorption and scattering of local light fields in the Fourier domain in the neighorhood of light paths. This analysis aims at predicting the changes over the distribution of light energy, so as to allow efficient sampling and integration methods of diffused light in participating media. Our analysis explains that absorbtion increases frequency since it acts as a continuous visibility mask over the local light field, and that scattering lowers frequencies as it operates a low pass convolution filter in the directional domain. In order to combine this analysis with our previous work on covariance tracing—and therefore use it to improve existing algorithms for path tracing in participating media—we derived new sampling metrics all based on a common prediction of the 3D covariance of the fluence in the volume. We demonstrate indeed that the covariance of the fluence can efficiently be computed by combining the 4D covariance matrices of light fields in the neighborhood of light paths, and that it can be used to compute effective metrics (1) for the variance of energy collected along camera rays, (2) for determining the shape and size of reconstruction kernels in screen space, and (3) for drastically improving the convergence of density estimation methods. For the later, we propose an improvement of the method of Progressive Photon Beams. This work has been published in ACM Transactions on Graphics and presented at Siggraph'2014 in Vancouver [5].

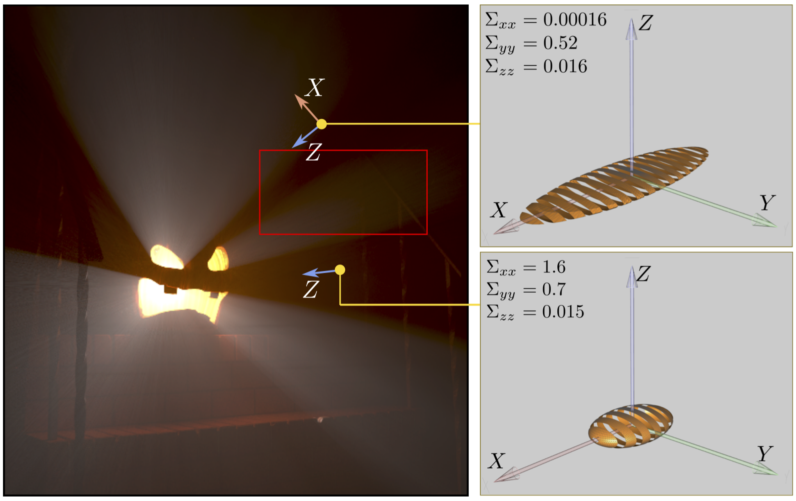


Figure 14. Predictions of the covariance of the Fourier spectrum of the fluence in the volume computed using our Fourier analysis of scattering and absorption.

Maxplus Project-Team

6. New Results

6.1. Highlights of the Year

Nous avons donné un contre exemple inattendu à l'analogue continu de la conjecture de Hirsch, proposé par Deza, Terlaky et Zinchenko, voir Section 6.4.4 .

English version

We gave a somehow unexpected counter example to the continuous analogue of the Hirsch conjecture proposed by Deza, Terlaky and Zinchenko, see Section 6.4.4 .

6.2. Théorie spectrale max-plus et géométrie métrique/Max-plus spectral theory and metric geometry

6.2.1. Introduction

Participants: Marianne Akian, Stéphane Gaubert, Cormac Walsh.

Étant donné un noyau $a : S \times S \to \mathbb{R} \cup \{-\infty\}$, on peut lui associer le problème spectral max-plus

$$\sup_{y \in S} a(x,y) + u(y) = \lambda + u(x), \quad \forall x \in S, \tag{8}$$

dans lequel on cherche le vecteur propre $u : S \to \mathbb{R} \cup \{-\infty\}$ et la valeur propre correspondante $\lambda \in \mathbb{R} \cup \{-\infty\}$. Comme nous l'avons rappelé dans les §3.2 et 3.3 , le problème spectral (9) intervient en contrôle ergodique: l'ensemble S est l'espace des états, et l'application $a(x,y)$ fournit le gain associé à la transition $x \to y$. Le cas où S est fini est classique, l'on a alors un résultat précis de représentation de l'espace propre, à l'aide d'un certain graphe, dit graphe critique. Des résultats existent également lorsque S est compact et que le noyau vérifie certaines propriétés de régularité.

Dans [64], nous avons considéré le cas où S est non compact. Lorsque $\lambda = 0$, l'espace propre est analogue à l'espace des fonctions harmoniques défini en théorie (classique ou probabiliste) du potentiel. En introduisant l'analogue max-plus de la frontière de Martin, nous avons obtenu un analogue de la formule de représentation de Poisson des fonctions harmoniques : toute solution u de (9) peut être représentée sous la forme :

$$u = \sup_{w \in \mathcal{M}_m} w + \mu_u(w) \quad , \tag{9}$$

où $\mathcal{M}_m \subset (\mathbb{R} \cup \{-\infty\})^S$ est l'analogue max-plus de la frontière de Martin minimale (l'ensemble des fonctions harmoniques extrémales normalisées), et où μ_u joue le rôle de la mesure spectrale. Nous avons montré aussi que les éléments de l'espace de Martin minimal peuvent être caractérisés comme les limites de “quasi-géodésiques”. La frontière de Martin max-plus généralise dans une certaine mesure la frontière d'un espace métrique construite à partir des horo-fonctions (fonctions de Busemann généralisées), ou horo-frontière. Ces résultats inspirent les travaux des sections suivantes, qui portent sur des cas remarquables d'espaces métriques (§6.2.3) ou sur des applications en théorie des jeux (§6.2.2).

English version

Let the kernel $a : S \times S \to \mathbb{R} \cup \{-\infty\}$ be given. One may associate the max-plus spectral equation (9), where the eigenvector $u : S \to \mathbb{R} \cup \{-\infty\}$ and the eigenvalue $\lambda \in \mathbb{R} \cup \{-\infty\}$ are unknown. As we recalled in §3.2 and refmonotone, this spectral problem arises in ergodic optimal control: the set S is the *state space*, and the map $a(x, y)$ is the *transition reward*. The case when S is finite is classical, a precise spectral theorem is known, with a characterisation of the eigenspace in terms of a critical graph. Some results have been shown when S is compact, assuming that the kernel a satisfies some regularity properties.

In [64], we considered the case where S is non-compact. When $\lambda = 0$, the eigenspace is analoguous to the set of harmonic functions defined in classical or probabilistic potential theory. By introducing a max-plus analogue of the classical Martin boundary, we obtained an analogue of the Poisson representation of harmonic functions, showing that any solution u of (9) may be represented as in (10) where $\mathcal{M}_m \subset (\mathbb{R} \cup \{-\infty\})^S$ is a max-plus analogue of the minimal Martin boundary (the set of normalised extremal harmonic functions), and μ_u plays the role of the spectral measure. We also showed that the elements of the minimal Martin boundary can be characterised as limits of certain "almost-geodesics". The max-plus Martin boundary generalises to some extent the boundary of metric spaces defined in terms of horofunctions (generalised Busemann functions), or horoboundary. These results have inspired the work of the next sections, which deal either with interesting examples of metric spaces (§6.2.3) or with applications to zero-sum games (§6.2.2).

6.2.2. Asymptotiques d'itérées d'applications contractantes au sens large et jeux à somme nulle en horizon long/Asymptotics of iterates of nonexpansive mappings and zero-sum games

Participants: Jérôme Bolte, Stéphane Gaubert, Guillaume Vigeral.

On s'est intéressé ici à l'existence du paiement moyen pour les jeux répétés, et plus généralement, à l'existence du vecteur de "taux de fuite" $\lim_k f^k(x)/k$ où f est une application de $\mathbb{R}^n$ dans lui même, nonexpansive pour une norme quelconque. Dans le cas particulier des jeux, f est un opérateur de Shapley, qui est nonexpansif pour la norme sup. On a montré dans [15] que la limite existe si l'application f est définissable dans une structure o-minimale. Ceci généralise des résultats de Bewley, Kohlberg, et Neyman, qui montraient que la limite existe si f est semi-algébrique. L'extension au cas o-minimal permet notamment de traiter des opérateurs de type "log-exp" apparaissant en contrôle sensible au risque.

English version

We studied the question of the existence of the mean payoff for repeated games, and more generally, the existence of a vector of "escape rates", $\lim_k f^k(x)/k$, where f is a self-map of $\mathbb{R}^n$, non-expansive in some norm. In the special case of zero-sum games, f is a Shapley operator, and it is sup-norm nonexpansive. We showed in [15] that this limit does exist as soon as the map f is definable in an o-minimal structure. This generalizes results of Bewley, Kohlberg, and Neyman, who showed that this limit exists if f is semi-algebraic. The extension to the case of o-minimal structures allows one in particular to deal with log-exp type operators arising in risk sensitive control.

6.2.3. Isométries de la géométrie de Hilbert/Isometries of the Hilbert geometry

Participants: Cormac Walsh, Bas Lemmens [Kent University, UK].

Dans nos travaux précédents, nous avons étudié la géométrie de Hilbert (d'un ensemble convexe) en dimension finie, en particulier son horo-frontière et son groupe des isométries. Le chapitre de livre [44] donne une vue d'ensemble de ces travaux. Le cas de la dimension infinie est aussi intéressant, et a été utilisé depuis de nombreuses années en analyse non linéaire. Malgré cela, la géométrie de ces espaces est très peu connue en dimension infinie. Nous collaborons sur ce sujet avec Bas Lemmens de l'Université de Kent. Nous étudions par exemple le problème suivant. En dimension finie, il est connu que la géométrie de Hilbert est isométrique à un espace normé si et seulement si le convexe est un simplexe. Nous essayons de montrer plus généralement que la géométrie de Hilbert est isométrique à un espace de Banach si et seulement si le convexe est le cône des fonctions positives continues sur un espace topologique compact. Pour cela, nous étudions l'horo-frontière en dimension infinie.

English version

Previously, we have been studying the Hilbert geometry in finite dimensions, especially its horofunction boundary and isometry group. The book chapter [44] contains a survey of this work. However, the infinite dimensional case is also interesting, and has been used as a tool for many years in non-linear analysis. Despite this, very little is known about the geometry of these spaces when the dimension is infinite. We are collaborating on this topic with Bas Lemmens of the University of Kent. An example of a problem we are working on is the following. In finite dimension it is known that a Hilbert geometry is isometric to a normed space if and only if it is a simplex. We are attempting to show that, more generally, a Hilbert geometry is isometric to a Banach space if and only if it is the cross-section of a positive cone, that is, the cone of positive continuous functions on some compact topological space. To tackle this problem we are finding it useful to study the horofunction boundary in the infinite-dimensional case.

6.2.4. Croissance des boules dans la géométrie de Hilbert/Volume growth in the Hilbert geometry

Participants: Cormac Walsh, Constantin Vernicos [Université Montpellier 2].

Avec Constantin Vernicos de l'Université Montpellier 2, nous étudions la croissance du volume de la boule d'une géométrie de Hilbert (d'un ensemble convexe) en fonction du rayon. En particulier, nous étudions l'entropie volumique:

$$\lim_{r\to\infty} \frac{\log \mathrm{Vol}\, B(x,r)}{r}, \tag{10}$$

où $B(x,r)$ désigne la boule de centre x et de rayon r, et Vol est une notion de volume particulière, telle que celle définie par Holmes–Thompson ou celle de Busemann. L'entropie ne dépend pas du choix particulier de x, ni de celui du volume. Il est connu que pour l'espace hyperbolique, ou toute géométrie de Hilbert dont la frontière est C^2 et de courbure strictement positive, l'entropie est égale à $n-1$ lorsque la dimension de l'espace est n, et il a été conjecturé que ceci correspond aussi à l'entropie maximale d'une géométrie de Hilbert en dimension n. Afin de prouver cette conjecture, nous cherchons d'abord à étudier le lien entre l'entropie et l'approximabilité du convexe par des polytopes, et ensuite à borner cette approximabilité. La première étape nécessite d'étudier la croissance du volume dans le cas de polytopes. Dans ce cas, la croissance est polynomiale de degré n, plutôt qu'exponentielle, et il est important de comprendre le lien entre le coefficient dominant du polynôme exprimant le volume et la complexité du polytope. Nous avons obtenu une formule pour ce coefficient, laquelle dépend de la structure combinatoire du polytope.

English version

In a collaboration with Constantin Vernicos of Université Montpellier 2, we are investigating how the volume of a ball in a Hilbert geometry grows as its radius increases. Specifically, we are studing the volume entropy (11) where $B(x,r)$ is the ball with center x and radius r, and Vol denotes some notion of volume, for example, the Holmes–Thompson or Busemann definitions. Note that the entropy does not depend on the particular choice of x, nor on the choice of the volume. It is known that the hyperbolic space, or indeed any Hilbert geometry with a C^2-smooth boundary of stricty positive curvature, has entropy $n-1$, where n is the dimension, and it has been conjectured this is the maximal entropy possible for Hilbert geometries of the given dimension. Our approach to this conjecture is to first relate the entropy to the approximability of the convex domain by polytopes, and then bound this approximability. The first of these steps requires us study the volume growth in the polytopal case. Here the growth is polynomial rather than exponential, of degree n, and it is important to know how the constant on front of the highest term depends on the complexity of the polytope. We have a formula for this constant in terms of the combinatorial structure of the polytope.

6.2.5. Consensus non-commutatif et contraction d'opérateurs de Kraus/Noncommutative consensus and contraction of Kraus maps

Participants: Stéphane Gaubert, Zheng Qu.

Dans le travail [17], on s'est intéressé à la vitesse de convergence vers l'équilibre d'une itération de la forme $x^{k+1} = T(x^k)$, $x^k \in X$, où T est une application linéaire préservant un cône dans un espace de Banach X, telle que $T(e) = e$, pour un certain vecteur e dans l'interieur du cône. On s'intéresse aussi à l'itération dans l'espace dual, $y^{k+1} = T^*(y^k)$, $y^k \in X^*$, lorsque $\langle y^0, e\rangle = 1$.

Le cas classique est celui où $T(x) = Px$ est un opérateur de Markov. L'itération primale traduit alors la convergence vers le "consensus", et l'itération duale traduit la convergence de la distribution de probabilité en temps k vers l'état stationnaire. Dans ce cas, le taux de contraction (en un coup) $\kappa(P)$ d'une itération primale, pour la semi-norme de Hilbert $\|z\|_H := \max_i z_i - \min_j z_j$, ainsi que le taux de contraction d'une itération duale, pour la métrique en variation totale, coïncident et sont caractérisés par une formule dûe à Doeblin et Dobrushin (coefficient d'ergodicité),

$$\kappa(P) := 1 - \min_{i,j} \sum_{s=1}^{n} \min\left(P_{is}, P_{js}\right).$$

On a donné ici une généralisation de cette formule au cas d'opérateurs abstraits, qui s'applique en particulier aux opérateurs de Kraus qui interviennent en information quantique. Ces derniers opérent sur l'espace des matrices symmétriques, et sont de la forme

$$T(x) = \sum_k a_k x a_k^* \qquad \text{avec} \qquad \sum_k a_k a_k^* = I \ .$$

Dans [53], nous avons étudié des questions de complexité pour les applications de Kraus, montrant en particulier qu'il est NP-dur de vérifier qu'une application de Kraus envoie le cone dans son interieur.

English version

In [17], we studied the speed of convergence to equilibrium of an iteration of the form $x^{k+1} = T(x^k)$, $x^k \in X$, where T is a linear map preserving a cone in a Banach space X, such that $T(e) = e$, for some vector e in the interior of the cone. We also considered the iteration in the dual space X^*, $y^{k+1} = T^*(y^k)$, $y^k \in X^*$, where $\langle y^0, e\rangle = 1$.

The classical application arises when $T(x) = Px$ is a Markov operator. Then, the primal iteration represents the dynamics of consensus, whereas the dual iteration represents the evolution of the probability distribution as a function of time. Then, the (one-shot) contraction rate $\kappa(P)$ of the primal iteration, with respect to Hilbert's seminorm $\|z\|_H := \max_i z_i - \min_j z_j$, and the contraction rate of the dual iteration, with respect to the total variation metric, coincide, and are characterized by a formula of Doeblin and Dobrushin (ergodicity coefficient),

$$\kappa(P) := 1 - \min_{i,j} \sum_{s=1}^{n} \min\left(P_{is}, P_{js}\right).$$

We gave here a generalization of this formula to an abstract operators on a cone. This covers in particular the Kraus maps arising in quantum information theory. The latter maps act on the space of symmetric matrices. They can be written as

$$T(x) = \sum_k a_k x a_k^* \qquad \text{with} \qquad \sum_k a_k a_k^* = I \ .$$

In [53], we studied complexity issues related to Kraus maps, and showed in particular that checking whether a Kraus map sends the cone to its interior is NP-hard.

6.3. Algèbre linéaire max-plus, convexité tropicale et jeux à somme nulle/Max-plus linear algebra, tropical convity and zero-sum games

6.3.1. *Polyèdres tropicaux/Tropical polyhedra*

Participants: Xavier Allamigeon, Stéphane Gaubert, Eric Goubault [CEA], Ricardo Katz [Conicet, Argentine].

On étudie les analogues max-plus ou tropicaux des ensembles convexes. Ceux-ci sont utiles en particulier pour représenter de manière effective les ensembles d'états accessibles de systèmes à événements discrets [9], ils sont aussi apparus récemment en géométrie tropicale, dans toute une série de travaux à la suite de Sturmfels et Develin [108]. Les polyèdres max-plus peuvent aussi être vus comme des limites de déformations de polyèdres classiques, sur lesquels ils donnent un éclairage de nature combinatoire. Toutes ces motivations ont inspiré la recherche d'analogues des résultats fondamentaux d'analyse convexe classique: séparation, projection, points extrémaux, à la suite en particulier de [8].

Dans un travail de X. Allamigeon, S. Gaubert, et E. Goubault [71], [72], on a mis en évidence un critère combinatoire pour la caractérisation des sommets des polyèdres tropicalement convexes. Celui-ci s'exprime à l'aide d'hypergraphes orientés, et de leurs composantes fortement connexes. Ce critère possède la propriété d'être vérifiable en un temps presque linéaire en la taille de l'hypergraphe.

On en déduit un analogue tropical de la méthode de la double description [72] (méthode très utilisée sur les polyèdres classiques, et dûe à Motzkin *et al.* [160]). Cet algorithme permet de calculer les sommets d'un polyèdre défini de façon externe (intersection de demi-espaces ou d'hyperplans tropicaux). Grâce au critère combinatoire précédent, l'algorithme améliore de plusieurs ordres de grandeur les techniques connues jusqu'alors. Ceci est confirmé par de nombreuses expérimentations. Ce travail est motivé par des applications à l'analyse statique [70] et aux systèmes à événements discrets [113], dans lesquelles la manipulation de tels polyèdres est le goulot d'étranglement.

Il est connu qu'un polyèdre tropical peut être représenté comme l'enveloppe convexe d'un ensemble minimal de points et rayons, donnés par ses sommets et ses rayons extrêmes [124]. Dans un travail réalisé par X. Allamigeon et R. Katz [75], et effectué en partie lors de visites de R. Katz à Inria, on étudie la question duale de la caractérisation des représentations minimales par demi-espaces. On montre qu'un polyèdre tropical possède *essentiellement* une unique représentation minimale par demi-espaces, lorsque leurs apex appartiennent au polyèdre. On montre que les apex de ces demi-espaces non-redondants correspondent à certains sommets du complexe tropical introduit par Develin et Sturmfels [108]. On introduit également un critère combinatoire pour l'élimination de demi-espaces redondants à l'aide d'hypergraphes orientés.

Dans un travail de X. Allamigeon et R. Katz [52], nous étudions la tropicalisation des représentations par demi-espaces des polyèdres convexes sur le corps des séries de Puiseux. Nous démontrons ainsi une conjecture de Develin et Yu [109]. Celle-ci assure qu'étant donné un polytope tropical pur, il existe un polytope *relevé* sur les séries de Puiseux, dont les demi-espaces associés aux faces se "tropicalisent" en une représentation par demi-espaces du polytope tropical initial.

Des applications de ces travaux à l'algorithmique, concernant en particulier les jeux répétés, sont discutées dans la Section 6.5.2 . Une application aux systèmes temps réel est discutée dans la Section 6.6.4 .

English version

We study the max-plus or tropical analogues of convex sets. These have been used in particular to represent effectively the accessible sets of certain discrete event systems [9]. They also appeared in tropical geometry, following the work of Sturmfels and Develin [108]. Max-plus polyhedra can be thought of as limits of deformations of classical polyhedra, on which they give a combinatorial insight. These motivations have inspired the investigation of analogues of basic results of classical convex analysis: separation, projection, representation by extreme points, following [8].

In a work of X. Allamigeon, S. Gaubert, and E. Goubault [72], we introduce a combinatorial criterion for the characterization of the vertices of tropically convex polyhedra. It is expressed in terms of directed hypergraphs and their strongly connected components. This criterion can be verified in almost linear time in the size of the hypergraph.

This allows to develop a tropical analogue of the double description method [72] (this method is widely used for classical convex polyhedra, and is due to Motzkin *et al.* [160]). This algorithm is able to determine all the vertices of a polyhedron defined externally (intersection of tropical half-spaces of hyperplanes). Thanks to the combinatorial criterion mentioned above, the algorithm improves the existing methods by several orders of magnitude. This is confirmed by several experiments. This is motivated by applications to static analysis [70] and discrete event systems [113], in which computing such polyhedra turns out to be the bottleneck.

It is well-known that a tropical polyhedron can be represented as the convex hull of a minimal set of points and rays, provided by its vertices and extreme rays [124]. In a work of X. Allamigeon and R. Katz [75], partly done during visits of R. Katz at Inria, the dual problem of characterizing the minimal representations by half-spaces is studied. We show that a tropical polyhedron admits *essentially* a unique minimal external representation by half-spaces, provided that their apices belong to the polyhedron. We prove that the apices of these half-spaces correspond to certain vertices of the tropical complex introduced by Develin and Sturmfels [108]. We also establish a combinatorial criterion allowing to eliminate redundant half-spaces using directed hypergraphs.

In a work of X. Allamigeon and R. Katz [52], we study the tropicalization of the representation by half-spaces of convex polyhedra over the field of Puiseux series. In particular, we prove a conjecture of Develin and Yu [109]. It states that, given a pure tropical polytope, there exists a lifting polytope over Puiseux series, such that the facet-defining half-spaces are “tropicalized” into a representation by half-spaces of the initial polytope.

Some algorithmic applications of this work concerning in particular mean payoff games, will be discussed in Section 6.5.2 . Applications to real time systems will be discussed in Section 6.6.4 .

6.3.2. Systèmes linéaires max-plus/Max-plus linear systems

Participants: Marianne Akian, Stéphane Gaubert, Alexander Guterman [Moscow State University].

Dans [42], on montre des formules de Cramer pour des systèmes linéaires sur diverses extensions du semi-anneau max-plus. Les éléments de ces extensions sont des nombres tropicaux enrichis d’une information de multiplicité, de signe ou d’angle par exemple. On obtient ainsi des résultats d’existence et d’unicité qui généralisent plusieurs résultats de [133], [164], [120], [172], [139]. De plus, pour certaines extensions du semi-anneau max-plus, les preuves fournissent des algorithmes de type Jacobi ou Gauss-Seidel pour résoudre les systèmes linéaires.

Nous nous intéressons maintenant à la complexité de la solution de systèmes linéaires tropicaux signés, i.e. de systèmes sur l’extension du semi-anneau max-plus avec signes, ou d’hyperplans sur ce semi-anneau.

English version

In [42], we prove general Cramer type theorems for linear systems over various extensions of the tropical semiring, in which tropical numbers are enriched with an information of multiplicity, sign, or argument. We obtain existence or uniqueness results, which extend or refine earlier results in [133], [164], [120], [172], [139]. Moreover, some of our proofs lead to Jacobi and Gauss-Seidel type algorithms to solve linear systems in suitably extended tropical semirings.

We study now the complexity of the solution of signed tropical linear systems, that is systems on the extension of the tropical semiring with signs, or the one of the nonemptyness of hyperplanes over this semiring.

6.3.3. Convexes tropicaux et théorème de Choquet/Tropical convex sets and Choquet theorem

Participants: Marianne Akian, Stéphane Gaubert, Paul Poncet.

La thèse de Paul Poncet [165] concernait essentiellement ce que l'on appelle l'analyse idempotente, c'est-à dire l'étude des espaces fonctionnels ou linéaires de dimension infinie sur l'algèbre tropicale, ou tout autre semi-anneau idempotent. Paul Poncet a développé pour cela un point de vue treillis continus comme dans [1], ou plus généralement domaines. Depuis la soutenance, plusieurs articles issus du manuscrit de thèse sont publiés [21], [20] ou en cours de soumission, et d'autres travaux pousuivant ceux de la thèse sont en cours avec les membres de l'équipe.

En particulier avec ce point de vue domaines, Paul Poncet a pu établir des théorèmes de type Krein-Milman, réciproque de Milman, et représentation de Choquet dans les semi-treillis [20] ou l'algèbre max-plus [38].

On sait que les résultats sur les convexes tropicaux de dimension infinie de [165] permettent de retrouver partiellement les résultats sur la frontière de Martin max-plus décrits dans la section 6.2.1 . Dans un travail commun avec Klaus Keimel (TU-Darmstadt), nous essayons d'obtenir l'extension du théorème de représentation de Choquet tropical dans le cas d'ensembles ordonnés qui ne sont pas forcément des treillis tels que le cône des matrices symmetriques positives muni de l'ordre de Loewner.

English version

The PhD thesis work of Paul Poncet [165] concerned essentially what is called idempotent analysis, that is the study of infinite dimensional functional or linear spaces over tropical algebra, or any other idempotent semiring. For this aim, Paul Poncet developped the point of view of continuous lattices, as in [1], or more generally of domains. Since the defense of his thesis, several papers derived from the thesis manuscript have been published [21], [20] or up to be submitted. Some other works pursuing the thesis work are done with team members.

In particular, using the point of view of domains, Paul Poncet showed results such as a Krein-Milman type theorem, a Milman converse type theorem, and a Choquet representation type theorem over semilattices [20] or over max-plus algebra [38].

We know that the results on infinite dimensional tropical convex sets of [165] allow one to recover at least partially the results on max-plus Martin boundaries described in Section 6.2.1 . In a joint work with Klaus Keimel (TU-Darmstadt), we try to obtain the extension of the max-plus Choquet representation theorem in the case of ordered sets that are not necessarily semilattices, such as the cone of nonnegative symmetric matrices endowed with the Loewner order.

6.3.4. Points fixes d'applications monotones homogènes et jeux à somme nulle/Fixed points of order preserving homogeneous maps and zero-sum games

Participants: Marianne Akian, Stéphane Gaubert, Antoine Hochart.

Pour les jeux répétés à somme nulle, un problème de base est de savoir si le paiement moyen par unité de temps est indépendant de l'état initial. Ici, on définit le paiement moyen directement au moyen de l'opérateur de Shapley (ou de la programmation dynamique) du jeu, lequel préserve l'ordre et commute avec l'addition d'une constante. Dans le cas particulier des jeux à zero joueur, i.e. de chaînes de Markov avec fonctionnelle additive, la solution du problème ci-dessus est fournie par le théorème ergodique. Dans [46], on généralise ce résultat au cas des jeux répétés à espace d'états fini. Cette généralisation est basée sur l'étude de la sous-classe d'opérateurs de Shapley *sans-paiement* (le paiement a lieu seulement le dernier jour), lesquels commutent avec la multiplication par une constante positive. L'intérêt de cette sous-classe est qu'elle inclue la fonction de récession d'un opérateur de Shapley, lorsqu'elle existe. Nous montrons que le paiement moyen est indépendant de l'état initial pour toutes les perturbations des paiements instantannés dépendantes de l'état si, et seulement si, une condition d'ergodicité est vérifiée. Cette dernière est caractérisée par l'unicité, à constante additive près, du point fixe de la fonction de récession de l'opérateur de Shapley, ou, dans le cas particulier des jeux stochastiques à nombre fini d'actions et information parfaite, par une condition d'accessibilité dans un hypergraphe orienté, entre deux sous-ensembles conjugués d'états. On montre aussi que l'ergodicité d'un jeu ne dépend que de la probabilité de transition et qu'elle peut être vérifiée en temps polynomial lorsque le nombre d'états est fixé.

Lorsque un jeu est ergodique au sens ci-dessus, son paiement moyen est indépendant de l'état initial, et il coincide avec la valeur propre non linéaire de l'opérateur de Shapley. De plus, le vecteur propre associé, appelé biais, permet de déterminer les stratégies stationnaires optimales. Un autre problème est alors de comprendre pour quelles classes de jeux, le biais est unique (à constante additive près). Dans [25], on considère des jeux avec un nombre fini d'états et d'actions, de paiements instantannés variables, mais de probabilités de transition fixées. On montre que le vecteur de biais, considéré comme une fonction des paiements instantannés, est unique génériquement (à constante additive près).

English version

A basic question for zero-sum repeated games consists in determining whether the mean payoff per time unit is independent of the initial state. Here the mean payoff is defined in terms of the Shapley operator (dynamic programming operator) of the game, which is an order preserving map commuting with the addition of a constant. In the special case of "zero-player" games, i.e., of Markov chains equipped with additive functionals, the answer to the above question is provided by the mean ergodic theorem. In [46], we generalize this result to repeated games with a finite state space. This generalization is based on the study of the subclass of *payment-free* Shapley operators (the payment only occurs when the game stops), which are commuting with the multiplication by a positive constant, and which include the recession function of any Shapley operator, when it exists. We show that the mean payoff is independent of the initial state for all state-dependent perturbations of the rewards if and only if an ergodicity condition is satisfied. The latter is characterized by the uniqueness modulo additive constants of the fixed point of the recession function of the Shapley operator, or, in the special case of stochastic games with finite action spaces and perfect information, by a reachability condition involving conjugate subsets of states in directed hypergraphs. We show that the ergodicity condition for games only depends on the support of the transition probability and that it can be checked in polynomial time when the number of states is fixed.

Under the above ergodicity condition, the mean payoff of the game is independent of the initial state, and it is characterized as the nonlinear eigenvalue of the Shapley operator. Moreover, the associated eigenvector, also called the bias, allows one to determine optimal stationary strategies. Then, another basic question is to understand for which classes of games the bias vector is unique (up to an additive constant). In [25], we consider games with finite state and action spaces, thinking of the transition payments as variable parameters, transition probabilities being fixed. We show that the bias vector, thought of as a function of the transition payments, is generically unique (up to an additive constant).

6.4. Algèbre max-plus, déformations et asymptotiques /Max-plus algebra, deformations and asymptotic analysis

6.4.1. Introduction

Comme indiqué dans le §3.7 , l'algèbre max-plus est la limite d'une déformation de l'algèbre classique, ou plutôt du semi-corps des réels positifs. Elle peut aussi fournir des estimations de ces déformations, puisque

$$\max(a,b) \leq \epsilon \log(e^{a/\epsilon} + e^{b/\epsilon}) \leq \epsilon \log(2) + \max(a,b) \quad . \tag{11}$$

L'utilisation de ces propriétés a déjà conduit dans le passé aux travaux sur les perturbations de valeurs propres [57], [56], [55], ou sur les grandes déviations [1], [61]. Dans les travaux qui suivent, nous exploitons ces propriétés dans des contextes reliés ou similaires à ceux de nos travaux précédents.

English version

As detailled in §3.7 , max-plus algebra is the limit of a deformation of classical algebra, or more precisely of the semi-field of usual real positive numbers. It can also give estimations for these deformations using for instance (12). By using these properties, we already obtained some works on singular perturbations of matrix eigenvalues [57], [56], [55], or on large deviations [1], [61]. In the works described below, we are exploiting again these properties in contexts that are related or similar to those of our earlier works.

6.4.2. Méthodes tropicales de localisation de valeurs propres de matrices/Tropical methods for the localisation of matrix eigenvalues

Participants: Marianne Akian, Stéphane Gaubert, Andrea Marchesini.

Dans un travail avec Meisam Sharify [63], on a comparé les modules des valeurs propres d'un polynôme matriciel au moyen des racines tropicales du polynôme obtenu en appliquant une norme donnée aux coefficients. En particulier, on a obtenu des inégalités de type majorisation qui généralisent les bornes obtenues par Polya et Ostrowski dans le cas de polynômes scalaires.

Dans [12], on montre des inégalités de type majorisation entre les modules des valeurs propres d'une matrice et les valeurs propres tropicales de la matrice de ses modules. En particulier, les majorations généralisent l'inégalité de Friedland [119] concernant le rayon spectral.

Nous avons amélioré et généralisé ces inégalités [37], en appliquant différents changements de variables diagonaux à la matrice complexe initiale, lesquels sont construits à partir des variables duales du problème d'affectation optimale paramètrique construit à partir d'une matrice tropicale associée à la matrice complexe. En particulier, lorsqu'on les applique à une matrice companion par blocs, ces inégalités sont similaires à celles de [63].

English version

In a work with Meisam Sharify [63], we compared the moduli of the eigenvalues of a matrix polynomial to the tropical roots of a polynomial obtained by applying a norm to the coefficients of the original matrix polynomial. In particular, we obtained majorization type inequalities which generalize the bounds of Polya and Ostrowski available for scalar polynomials.

In [12], we show majorization type inequalities between the moduli of the eigenvalues of a complex matrix and the tropical eigenvalues of the matrix obtained by applying the modulus entrywise. In particular, the upper bounds generalize the inequality of Friedland [119] concerning the spectral radius. The above inequalities were obtained by using the permanental and tropical analogues of the exterior power of a matrix and by showing (combinatorially) properties of their eigenvalues similar to the ones of usual exterior powers.

We recently improved and generalized these inequalities, see [37], by applying to the original complex matrix, different diagonal scalings constructed from the dual variables of the parametric optimal assignment constructed from an associated tropical matrix. In particular, when applied to a block companion matrix, our inequalities are similar to the ones in [63].

6.4.3. Méthodes tropicales pour le calcul numérique de valeurs propres de matrices/Tropical methods for the numerical computation of matrix eigenvalues

Participants: Marianne Akian, Stéphane Gaubert, Andrea Marchesini.

Un des buts de la thèse d'Andrea Marchesini est d'utiliser les résultats de localisation de valeurs propres tels que ceux obtenus ci-dessus pour améliorer la précision des algorithmes de calcul numérique de valeurs propres de matrices ou de polynômes matriciels, en particulier en construisant des changements d'échelle exploitant les calculs tropicaux, à effectuer préalablement à l'appel d'algorithmes classiques comme QZ. Le "changement d'échelle tropical" introduit par Stéphane Gaubert et Meisam Sharify [127] dans le cas de polynôme matriciels quadratiques consiste en un changement de variable multiplicatif de la variable scalaire du polynôme matriciel. Dans un travail en collaboration avec Françoise Tisseur et James Hook de l'Université de Manchester [36], on considère aussi un changement de variables diagonal du polynôme matriciel construit à partir des variables duales du problème d'affectation optimale paramètrique construit dans l'esprit de [55]. On montre l'intérêt de ces changements d'échelle en terme de conditionnement des valeurs propres, et la supériorité du changement de variables diagonal par rapport au changement d'échelle tropical.

English version

One of goals of the PhD thesis of Andrea Marchesini is to use results on the localisation of eigenvalues like the above ones, to improve the accuracy of the numerical compution of the eigenvalues of a complex matrix or matrix polynomial, in particular by applying scaling methods using tropical techniques, which may be used before calling usual algorithms as QZ. The "tropical scaling" introduced by Stéphane Gaubert and Meisam Sharify [127] in the case of quadratic matrix polynomials consists in a multiplicative scaling of the scalar variable of the matrix polynomial. In a work with Françoise Tisseur and James Hook from Manchester University [36], we also consider a diagonal scaling of the matrix polynomial constructed from the dual variables of the parametric optimal assignment constructed in the same spirit as in [55]. We show the interest of these scaling methods on the eigenvalue conditioning, and the superiority of the diagonal scaling with respect to the tropical scaling.

6.4.4. Tropicalisation du chemin central, et application à la courbure/Tropicalization of the central path and application to the curvature

Participants: Xavier Allamigeon, Pascal Benchimol, Stéphane Gaubert, Michael Joswig [TU Berlin].

En optimisation, une classe importante d'algorithmes, dits *de points intérieurs*, consiste à suivre une courbe appelée *chemin central* jusqu'à atteindre la solution optimale. Le chemin central d'un programme linéaire $\mathrm{LP}(A,b,c) \equiv \min\{c \cdot x \mid Ax \leq b,\ x \geq 0\}$ est défini comme l'ensemble des solutions optimales (x^μ, w^μ) des problèmes à barrière logarithmique:

$$\begin{aligned} \text{minimiser} \quad & c \cdot x - \mu\Big(\sum_{j=1}^{n} \log x_j + \sum_{i=1}^{m} \log w_i\Big) \\ \text{sous les contraintes} \quad & Ax + w = b,\ \ x > 0,\ \ w > 0 \end{aligned}$$

Les performances d'un algorithme de point intérieur sont intimement liées à la forme du chemin central. En particulier, la courbure mesure de combien un chemin diffère d'une ligne droite. Intuitivement, un chemin central à forte courbure devrait être plus difficile à approximer par des segments de droites, ce qui suggère davantage d'itérations des algorithmes de points intérieurs. La courbure totale du chemin central a été étudiée par Dedieu, Malajovich et Shub [105] à travers le théorème de Bezout dans le cas multihomogène, et par De Loera, Sturmfels and Vinzant [104] à l'aide de la théorie des matroïdes. Ces deux travaux fournissent une borne supérieure en $O(n)$ sur la courbure totale moyenne sur l'ensemble des régions formées par l'arrangement d'hyperplans en dimension n. Le cube de Klee-Minty redondant de [111] et le "serpent" de [110] sont des instances qui montrent que la courbure totale peut être de l'ordre de $\Omega(m)$ pour un polytope défini par m inégalités.

Dans un travail de X. Allamigeon, P. Benchimol, S. Gaubert, and M. Joswig, nous avons étudié la tropicalisation du chemin central. Le *chemin central tropical* est défini comme la limite logarithmique des chemins centraux d'une famille paramétrique de programmes linéaires $\mathrm{LP}(A(t), b(t), c(t))$, où les entrées $A_{ij}(t)$, $b_i(t)$ and $c_j(t)$ sont des fonctions définissables dans une structure o-minimale appelée *corps de Hardy*.

Une première contribution a été de fournir une caractérisation entièrement géométrique du chemin central tropical. Nous avons montré que le centre analytique est donné par le plus grand élément de l'ensemble des points tropicaux admissibles. De plus, tout point du chemin central tropical coincide avec le plus grand élément de l'ensemble admissible tropical intersecté avec un ensemble de sous-niveau de la fonction de coût tropicale.

Grâce à cette caractérisation, nous avons réfuté l'analogue continu de la conjecture de Hirsch proposé par Deza, Terlaky et Zinchenko. l'analogue continu de la conjecture de Hirsch proposé par Deza, Terlaky et Zinchenko. Ainsi, nous avons construit une famille de programmes linéaires définis par $3r+4$ inequalities in dimension $2r+2$, où le chemin central a une courbure totale en $\Omega(2^r/r)$. Cette famille est obtenue en relevant des programmes linéaires tropicaux introduits par Bezem, Nieuwenhuis et Rodríguez-Carbonell [83] pour montrer qu'un algorithme de Butkovič and Zimmermann [88] a une complexité exponentielle. Leur chemin central tropical a une forme de courbe en escalier avec $\Omega(2^r)$ marches.

Ces résultats sont rassemblés dans le document [50]. Ils ont été présentés à la conférence [41].

English version

In optimization, path-following interior point methods are driven to an optimal solution along a trajectory called the central path. The *central path* of a linear program $\mathrm{LP}(A, b, c) \equiv \min\{c \cdot x \mid Ax \leq b,\ x \geq 0\}$ is defined as the set of the optimal solutions (x^μ, w^μ) of the barrier problems:

$$\begin{array}{ll} \text{minimize} & c \cdot x - \mu(\sum_{j=1}^{n} \log x_j + \sum_{i=1}^{m} \log w_i) \\ \text{subject to} & Ax + w = b,\ \ x > 0,\ \ w > 0 \end{array}$$

The performance of an interior point method is tightly linked to the shape of its central path. In particular, the curvature measures how far a path differs from a straight line. Intuitively, a central path with high curvature should be harder to approximate with line segments, and thus this suggests more iterations of the interior point methods. The total curvature of the central path has been studied by Dedieu, Malajovich and Shub [105] via the multihomogeneous Bézout Theorem and by De Loera, Sturmfels and Vinzant [104] using matroid theory. These two papers provide an upper bound of $O(n)$ on the total curvature averaged over all regions of an arrangement of hyperplanes in dimension n. The redundant Klee-Minty cube of [111] and the “snake” in [110] are instances which show that that the total curvature can be in $\Omega(m)$ for a polytope described by m inequalities. By analogy with the classical Hirsch conjecture, Deza, Terlaky and Zichencko [110] conjectured that $O(m)$ is also an upper bound for the total curvature.

In a work of X. Allamigeon, P. Benchimol, S. Gaubert, and M. Joswig, we have studied the tropicalization of the central path. The *tropical central path* is defined as the logarithmic limit of the central paths of a parametric family of linear programs $\mathrm{LP}(A(t), b(t), c(t))$, where the entries $A_{ij}(t)$, $b_i(t)$ and $c_j(t)$ are definable functions in an o-minimal structure called the *Hardy field*.

A first contribution is to provide a purely geometric characterization of the tropical central path. We have shown that the tropical analytic center is the greatest element of the tropical feasible set. Moreover, any point of the tropical central path is the greatest element of the tropical feasible set intersected with a sublevel set of the tropical objective function.

Thanks to this characterization, we disprove the continuous analog of the Hirsch conjecture proposed by Deza, Terlaky and Zinchenko, by constructing a family of linear programs with $3r + 4$ inequalities in dimension $2r + 2$ where the central path has a total curvature in $\Omega(2^r/r)$. This family arises by lifting tropical linear programs introduced by Bezem, Nieuwenhuis and Rodríguez-Carbonell [83] to show that an algorithm of Butkovič and Zimmermann [88] has exponential running time. The tropical central path looks like a staircase shape with $\Omega(2^r)$ steps.

These results are gathered in the preprint [50]. They have been presented in the conference [41].

6.5. Algorithmes/Algorithms

6.5.1. *Itération sur les politiques pour le contrôle stochastique et les jeux répétés à somme nulle/Policy iterations for stochastic control and repeated zero sum games*

Participants: Marianne Akian, Stéphane Gaubert.

L’algorithme d’itération sur les politiques est bien connu pour résoudre efficacement les équations de la programmation dynamique associées à des problèmes de contrôle stochastique avec critère à horizon infini (Howard) ou ergodique (Howard, et Denardo et Fox). Il a aussi été développé dans le cas de jeux à deux joueurs et somme nulle actualisés (Denardo) ou ergodiques (Hoffman et Karp).

Des résultats récents de Ye ainsi que Hansen, Miltersen et Zwick montrent que l'algorithme d'itération sur les politiques, restreint à la classe des jeux à somme nulle (à 1 ou 2 joueurs) actualisés de facteur d'actualisation donné, est fortement polynomial. Dans [58], on montre que ceci est le cas aussi pour l'algorithme d'itération sur les politiques pour les jeux à somme nulle et paiement moyen, restreint à la classe des jeux qui ont un temps moyen de retour ou d'arrivée à un état donné borné. La preuve utilise des techniques de théorie de Perron-Frobenius non-linéaire, permettant de ramener le problème à paiement moyen à un problème actualisé (de facteur d'actualisation dépendant de l'état et des actions). La même technique permet aussi de traiter le cas de jeux à somme nulle actualisés dont le facteur d'actualisation peut dépendre de l'état et des actions et prendre éventuellement des valeurs supérieures à 1. Récemment, on a montré que la borne pour le cas des jeux à somme nulle et paiement moyen s'applique aussi au cas des jeux actualisés de facteur d'actualisation constant [31], [32], [45]. Ce dernier résultat est inspiré par des résultats récents de Post et Ye et de Scherrer concernant les algorithmes du simplexe et d'itération sur les politiques pour les problèmes de contrôle optimal (ou jeux à 1 joueur).

English version

Policy iteration is a powerful and well known algorithm to solve the dynamic programming equation associated to stochatic control (one player game) problems with infinite horizon criterion (Howard) or ergodic criterion (Howard and Denardo and Fox). It has also be developped in the case of zero-sum two player games, either in discounted case (Denardo) or the ergodic one (Hoffman et Karp).

Recent results of Ye and Hansen, Miltersen and Zwick show that policy iteration for one or two player (perfect information) zero-sum stochastic games, restricted to instances with a fixed discount rate, is strongly polynomial. In [58], we show that policy iteration for mean-payoff zero-sum stochastic games is also strongly polynomial when restricted to instances with bounded first mean return time to a given state. The proof is based on methods of nonlinear Perron-Frobenius theory, allowing us to reduce the mean-payoff problem to a discounted problem with state dependent discount rate. Our analysis also shows that policy iteration remains strongly polynomial for discounted problems in which the discount rate can be state dependent (and even negative) at certain states, provided that the spectral radii of the nonnegative matrices associated to all strategies are bounded from above by a fixed constant strictly less than 1. Recently, we have proved that the bound for the case of mean-payoff zero-sum stochastic two-player games also holds for discounted games with a constant discount factor [31], [32], [45]. The latter result was inspired by recent results of Post and Ye, and Scherrer, concerning simplex and policy iteration algorithms for Markov decision processes (1 player games).

6.5.2. Algorithmique des polyèdres tropicaux/Algorithmics of tropical polyhedra

Participants: Xavier Allamigeon, Pascal Benchimol, Stéphane Gaubert, Eric Goubault [CEA], Michael Joswig [TU Berlin].

X. Allamigeon, S. Gaubert, et E. Goubault, ont développé dans [70], [72] plusieurs algorithmes permettant de manipuler des polyèdres tropicaux. Ceux-ci correspondent aux travaux décrits dans §6.3.1 . Ils permettent notamment de déterminer les sommets et rayons extrêmes d'un polyèdre tropical défini comme intersection de demi-espaces, ou inversement, de calculer une représentation externe à partir d'un ensemble de générateurs. Ces algorithmes sont implémentés la bibliothèque TPLib (voir §5.3).

Dans un travail de X. Allamigeon, P. Benchimol, S. Gaubert et M. Joswig [51], nous avons défini un analogue tropical de l'algorithme du simplexe qui permet de résoudre les problèmes de *programmation linéaire tropicale*, *i.e.*

$$\begin{array}{rl} \text{minimiser} & \max\limits_{1\le j\le n} c_j + x_j \\ \text{sous les contraintes} & \max\big(\max\limits_{1\le j\le n}(a^+_{ij}+x_j), b^+_i\big) \geq \max\big(\max\limits_{1\le j\le n}(a^-_{ij}+x_j), b^-_i\big), \quad i=1,\cdots,m \\ & x \in (\mathbb{R}\cup\{-\infty\})^n \end{array} \tag{12}$$

où les entrées du programme $a^{\pm}_{ij}$, $b^{\pm}_i$, c_j sont à valeur dans $\mathbb{R} \cup \{-\infty\}$. Ces problèmes sont intimement liés à la résolution de jeux répétés à somme nulle, puisque résoudre un jeux à paiement moyen déterministe est équivalent à déterminer si un problème de programmation linéaire admet un point réalisable [59].

Comme son homologue usuel, le simplexe tropical pivote entre des points de base (tropicaux), jusqu'à atteindre l'optimum du programme linéaire. La différence fondamentale avec l'algorithme du simplexe classique est que le pivotage est réalisé de manière purement combinatoire, en s'appuyant sur des descriptions locales du polyèdre tropical défini par les contraintes à l'aide d'(hyper)graphes orientés. Ceci nous a permis de prouver que *l'étape de pivotage (incluant le calcul des coûts réduits) a la même complexité en temps que dans l'algorithme classique, i.e.* $O(n(m+n))$. Ceci est d'autant plus inattendu que la structure des arêtes tropicales entre deux points de base sont géométriquement plus complexes (elles sont constituées de plusieurs segments de droite, jusqu'à n).

Le simplexe tropical a la propriété d'être fortement corrélé avec l'algorithme du simplexe classique. Grâce au principe de Tarski, le simplexe usuel peut être transposé tel quel sur des programmes linéaires dont les coefficients en entrée sont non plus des réels, mais sur le corps $\mathbb{R}\{\{t\}\}$ des séries de Puiseux généralisées en une certaine indéterminée t, *i.e.* des objets de la forme :

$$c_{\alpha_1} t^{\alpha_1} + c_{\alpha_2} t^{\alpha_2} + \cdots \tag{13}$$

où les α_i sont des réels, les coefficients c_{α_i} sont des réels non-nuls, et où la séquence des $\alpha_1, \alpha_2, \cdots$ est strictement croissante et soit finie, soit non-bornée. L'opposé du plus petit exposant de la série, $-\alpha_1$, est appelé *valuation* de la série. Un programme linéaire tropical est dit *relevé* en un problème linéaire sur $\mathbb{R}\{\{t\}\}$, si la valuation des coefficients en entrée de ce dernier sont égaux aux coefficients du problème tropical. Dans nos travaux, nous avons établi la correspondance suivante entre le simplexe usuel et le simplexe tropical : *pour tout programme linéaire tropical générique, l'algorithme du simplexe tropical trace l'image par la valuation du chemin sur l'algorithme du simplexe usuel sur n'importe quel relèvement du programme tropical dans* $\mathbb{R}\{\{t\}\}$.

Les résultats présentés ci-dessus sont rassemblés dans l'article [51]. Ils ont fait l'objet de plusieurs présentations en conférence [67], [68][27].

Ces résultats ouvrent la possibilité de relier la complexité du l'algorithme du simplexe usuel avec celles des jeux déterministes. Pour ces derniers, on sait seulement que leur résolution est dans la classe de complexité $\mathsf{NP} \cap \mathsf{coNP}$, et on ignore s'il existe un algorithme de complexité polynomiale. De façon similaire, on ne sait pas caractériser de façon précise la complexité de l'algorithme du simplexe usuel. Celle-ci dépend fortement de la règle de pivotage utilisée, et il existe des problèmes sur lesquelles de nombreuses règles de pivotage ont une complexité exponentielle. L'existence d'une règle de pivotage qui permettrait au simplexe de terminer en temps polynomial sur n'importe quelle instance est encore aujourd'hui une question ouverte.

Dans un deuxième travail, nous avons relié les deux problèmes ouverts précédents, grâce à l'algorithme du simplexe tropical. Nous avons en effet exhibé une classe de règles de pivotage, dites *combinatoires*, et avons montré qu'elles satisfont la propriété suivante : *s'il existe une règle de pivotage combinatoire qui permet de résoudre tout problème de programmation linéaire usuel en temps polynomial, alors on peut résoudre les jeux à paiement moyen en temps (fortement) polynomial.* Le terme *combinatoire* fait référence au fait que la règle est définie en fonction du signe des mineurs de la matrice des coefficients du problème linéaire. Ce résultat est décrit dans l'article [49], et a été présenté dans plusieurs conférences [39], [40].

Enfin, dans un travail de X. Allamigeon, P. Benchimol et S. Gaubert [26], nous avons étendu les résultats aux règles de pivotage *semi-algébriques*, classe incluant la règle dite du *shadow-vertex*. Celle-ci est connue pour avoir fourni plusieurs bornes de complexité moyenne et lisse sur l'algorithme du simplexe. Nous avons donc tropicalisé l'algorithme du simplexe shadow-vertex, et nous avons montré que cet algorithme permet de résoudre les jeux à paiement moyen en temps polynomial en moyenne.

English version

X. Allamigeon, S. Gaubert, and E. Goubault, have developed in [70], [72] algorithms allowing one to manipulate tropical polyhedra. They correspond to the contributions described in §6.3.1 . In particular, they can be used to determine the vertices and extreme rays of a tropical polyhedron defined as the intersection of half-spaces, or inversely, to compute an external description from a set of generators. These algorithms are implemented in the library TPLib (see §5.3).

In an ongoing work of X. Allamigeon, P. Benchimol, S. Gaubert and M. Joswig, we introduced a tropical analogue of the simplex algorithm, allowing one to solve problems of *tropical linear programming*, which are of the form (13), where the coefficients of the program, $a_{ij}^{\pm}$, $b_i^{\pm}$, c_j take their values in the max-plus semiring $\mathbb{R} \cup \{-\infty\}$. These problems are closely related to mean payoff games, as solving a game of this kind is equivalent to determine whether a tropical linear program admits a feasible point [59].

Like the classical simplex algorithm, the tropical simplex algorithm performs pivoting operations between basis points, until it reaches the optimum. The main discrepancy with the classical algorithm is that the pivoting is now a purely combinatorial operation, which is performed by using a local description of the polyhedron by a directed hypergraph. This allowed us to show that *a tropical pivoting step (including computing reduced costs) has the same complexity as in the classical simplex algorithm, i.e.* $O(n(m+n))$. This is all the more surprising as the tropical edge between two given points has a geometrically more complex structure in the tropical case (it is constituted of up to n ordinary line segments).

The tropical simplex algorithm turns out to be closely related to the classical one. Thanks to Tarski's principle, the latter is also valid for linear programs over the field $\mathbb{R}\{\{t\}\}$ of generalized Puiseux series in an indeterminate t. These series are of the form (14), where the α_i are real numbers, the coefficients c_{α_i} are non-zero reals, and the sequence $\alpha_1, \alpha_2, \cdots$ is strictly increasing and either finite or unbounded. The opposite of the smallest exponent of the series, $-\alpha_1$, is called *valuation*. A tropical linear program is said to be *lifted* to a linear program over $\mathbb{R}\{\{t\}\}$ if the valuation of the coefficients of the latter are sent to the coefficients of the former by the valuation. We showed the following relation between the classical simplex algorithm and its tropical analogue: *for all generic tropical linear program, the tropical simplex algorithm computes the image by the valuation of the path of the classical simplex algorithm, applied to any lift in* $\mathbb{R}\{\{t\}\}$ *of the original program.*

These results are gathered in the article [51]. They have been presented in several conferences [67], [68][27].

They allow one to relate the complexity of the classical simplex algorithm with the complexity of mean payoff games. The latter is unsettled, these games are known to be in the class NP $\cap$ coNP but it is not known whether they can be solved in polynomial time. Basic complexity issues regarding the classical simplex algorithm are also unsettled: its execution time depends on the pivoting rule, and many pivoting rules have been shown to have exponential worst case behaviors. The existence of a pivoting rule leading the simplex to terminate in polynomial time is still an open question. . In a second work, we related these two open questions, via the tropical simplex algorithm. We identified a class of pivoting rules, which are said to be *combinatorial*, and show that they have the following property: *if there is a combinatorial pivoting rule allowing one to solve every classical linear programming problem in polynomial time, then, mean payoff games can be solved in (strongly) polynomial time.* By *combinatorial*, we mean that the rule depends only of the coefficients of the system through the signs of minors of the coefficients matrix. This result is given in the article [49]. It has been presented to the conferences [39], [40].

Finally, in a work of X. Allamigeon, P. Benchimol and S. Gaubert [26], we extended the latter results to *semi-algebraic* pivoting rules, which include the so-called *shadow-vertex* rule. This rule has been exploited in the literature to establish several average-case and smooth complexity bounds on the simplex algorithm. We tropicalized the shadow-vertex simplex algorithm, and showed that it solves mean payoff games in polynomial time on average.

6.5.3. Problèmes d'accessibilité dans les hypergraphes orientés et leur complexité/Reachability problems in directed hypergraphs and their complexity

Participant: Xavier Allamigeon.

Les hypergraphes orientés sont une généralisation des graphes orientés, dans lesquelles chaque arc relie un ensemble de sommets à un autre. Ils jouent un rôle important dans les travaux récents sur la convexité tropicale (voir §6.3.1), puisqu'ils offrent une représentation naturelle des cônes définis sur le sous-semi-anneau booléen $\mathbb{B} = \{-\infty, 0\}$.

Dans un travail de X. Allamigeon [66], on étudie la complexité de problèmes d'accessibilité sur les hypergraphes orientés. Nous introduisons un algorithme de complexité presque linéaire permettant de déterminer les composantes fortement connexes terminales (qui n'accèdent à aucune autre composante si ce n'est elles-mêmes) d'un hypergraphe.

Nous établissons également une borne inférieure sur-linéaire sur la taille de la réduction transitive de la relation d'accessibilité dans les hypergraphes. Cela indique que la relation d'accessibilité dans les hypergraphes orientés est combinatoirement plus complexe que celle des graphes orientés. Cela suggère aussi que des problèmes comme le calcul des composantes fortement connexes est plus difficile sur les hypergraphes que sur les graphes. Nous mettons d'ailleurs en évidence une réduction en temps linéaire du problème du calcul des ensembles minimaux dans une famille d'ensembles donnée, vers le problème du calcul de toutes les composantes fortement connexes d'un hypergraphe. Le problème du calcul des ensembles minimaux a été largement étudié dans la littérature [166], [185], [184], [167], [168], [169], [115], [80], et aucune algorithme en temps linéaire n'est connu à ce jour.

English version

Directed hypergraphs are a generalization of directed graphs, in which the tail and the head of the arcs are sets of vertices. It appears that they play an important role in the recent works on tropical convexity (see §6.3.1), since they offer a natural representation of cones defined over the boolean sub-semiring $\mathbb{B} = \{-\infty, 0\}$.

In a work of X. Allamigeon [66], we study the complexity of reachability problems on directed hypergraphs. We introduce an almost linear-time algorithm allowing to determine the terminal strongly connected components (a component is said to be *terminal* when no other component is reachable from it).

We also establish a super-linear lower bound over the size of the transitive reduction of the reachability relation in directed hypergraphs. This indicates that the reachability relation is combinatorially more complex in directed hypergraphs than in directed graphs. This also suggests that reachability problems such as computing all strongly connected components are likely to be harder in hypergraphs than in graphs. Besides, we show that the minimal set problem can be reduced in linear time to the problem of computing all strongly connected components in hypergraphs. The former problem consists in finding all minimal sets among a given family of sets. It has been well studied in the literature [166], [185], [184], [167], [168], [169], [115], [80], and no linear time algorithm is known.

6.5.4. Approximation max-plus de fonctions valeurs et équations de Riccati généralisées/Max-plus approximation of value functions and generalized Riccati equations

Participants: Stéphane Gaubert, Zheng Qu.

Les méthodes d'approximation max-plus conduisent à approcher la fonction valeur d'un problème de contrôle ou de jeux par un supremum d'un nombre fini de formes quadratiques, voir notamment [126]. On s'intéresse ici à l'analyse théorique (complexité) ainsi qu'à l'amélioration de ces méthodes. Dans certains cas, ces formes quadratiques sont propagées par des flots d'équations de Riccati généralisées. Afin d'effectuer des analyses d'erreur, on exploite les propriétés de contraction du flot de Riccati pour certaines métriques connues sur le cône des matrices positives, et en particulier pour la métrique de Thompson. Celle-ci n'est rien d'autre que $d_T(A, B) = \| \log \mathrm{spec}\,(A^{-1}B)\|_\infty$, où spec désigne la suite des valeurs propres d'une matrice, et log s'entend composante par composante.

Ceci nous a amené à étudier le problème général du calcul du taux de contraction d'un flot monotone sur un cône, pour la métrique de Thompson. En effet, les propriétés de contraction de l'équation de Riccati standard sont connues (résultats de Bougerol pour la métrique Riemanienne invariante, et de Wojtowski pour la métrique de Thompson), mais les techniques de preuve employées dans ce cadre (semigroupes de matrices symplectiques) ne s'étendent pas aux équations généralisées.

On donne dans [16] une formule explicite générale pour le taux de contraction pour la métrique de Thompson d'un flot monotone, faisant seulement intervenir le générateur du flot et sa dérivée. On a notamment appliqué ce résultat à une équation de Riccati généralisée associé à des problèmes de contrôle stochastique avec critère quadratique, dans lesquels la dynamique comporte un terme bilinéaire en le contrôle et le bruit. On a montré dans ce cas que la métrique de Thompson est la seule métrique de Finsler invariante pour laquelle le flot est nonexpansif, et l'on a caractérisé la constante de contraction locale.

Une application de ces résultats de contraction à l'analyse d'une méthode de réduction de la malédiction de la dimension, dûe à McEneaney, a été donnée dans [22].

Une nouvelle méthode numérique maxplus, de nature randomisée, a été introduite dans [30], elle fait apparaître de très fortes accélérations par rapport aux méthodes précédentes.

La question de l'émondage des représentations max-plus a été abordée dans [29], où il est montré qu'une classe de relaxations convexes introduites par Sridharan et al. pour traiter numériquement un problème de contrôle quantique sont en fait exactes (pas de saut de relaxation).

English version

The max-plus methods lead to approach the value function of an optimal control or zero-sum game problem by a supremum of a finite number of quadratic forms, see in particular [126]. We are interested here in the theoretical analysis (complexity) of this class of methods, as well as of their improvement. In certain cases, the quadratic forms are propagated by the flows of generalized Riccati equations. In order to perform an error analysis, we need to use some contraction properties of the Riccati flow, for certain known metrics on the space of positive matrices, like Thompson's metric. The latter is nothing but $d_T(A,B) = \|\log\mathrm{spec}\,(A^{-1}B)\|_\infty$, where spec denotes the sequence of eigenvalues of a matrix, and $\log$ is understood entrywise.

This led us to study the general problem of computing the contraction rate of an order-preserving flow on a cone, with respect to Thompson's metric. Indeed, the contraction properties of the standard Riccati flow are known (theorem of Bougerol for the invariant Riemanian metric, of Wojtowski for the Thompson's metric), but the proof of these properties (based on symplectic semigroups) does not carry over to generalized Riccati equations.

We gave in [16] a general explicit formula for the contraction rate with respect to Thompson's metric of an order-preserving flow, involving only the generator of the flow and its derivative. We applied in particular this result to a generalized Riccati equation, associated to stochastic optimal control problems with a quadratic cost and a bilinear dynamics (presence of a bilinear term between the control and the noise). We showed that in this case, the Thompson's metric is the only invariant Finsler metric in which the generalized Riccati flow is nonexpansive, and we characterized the local contraction rate of this flow.

Z. Qu has applied these results in [22] to the analysis of a method of reduction of the curse of dimensionality, introduced by McEneaney.

A new max-plus numerical method, of a randomized nature, has been introduced in [30]. It shows an important speedup by by comparison with earlier methods.

The question of trimming max-plus representations was dealt with in [29]. It is shown there that a class of convex relaxations introduced Sridharan et al. to solve numerically some quantum control problem is exact.

6.5.5. Approximation probabiliste d'équations d'Hamilton-Jacobi-Bellman et itération sur les politiques

Participants: Marianne Akian, Eric Fodjo.

La thèse d'Eric Fodjo traite de problèmes de contrôle stochastique (de diffusions) avec critère à horizon infini actualisé ou arrêté, ou moyen en temps long, issus en particulier de problèmes de gestion de portefeuille avec coûts de transaction. La programmation dynamique conduit à une équation aux dérivées partielles d'Hamilton-Jacobi-Bellman, sur un espace de dimension au moins égale au nombre d'actifs risqués. La malédiction de la dimension ne permet pas de traiter numériquement ces équations en dimension grande (supérieure à 5). On se propose d'aborder ces problèmes avec des méthodes numériques associant itération sur les politiques, discrétisations probabilistes, et discrétisations max-plus, afin d'essayer de monter plus en dimension. Une autre piste est de remplacer l'itération sur les politiques par une approximation par des problèmes avec commutations optimales. Ces méthodes devraient aussi s'appliquer au cas de problèmes à horizon fini.

English version

The PhD thesis of Eric Fodjo concerns stochastic control problems with long term discounted or stopped payoff, or with mean-payoff in time, obtained in particular in the modelisation of portfolio selection with transaction costs. The dynamic programming method leads to a Hamilton-Jacobi-Bellman partial differential equation, on a space with a dimension at least equal to the number of risky assets. Curse of dimensionality does not allow one to solve numerically these equations for a large dimension (greater to 5). We propose to tackle these problems with numerical methods combining policy iterations, probabilistic discretisations, max-plus discretisations, in order to increase the possible dimension. Another solution is to replace policy iterations by an approximation with optimal switching problems. These methods should also be useful for finite horizon problems.

6.6. Applications

6.6.1. Introduction

Nous présentons maintenant plusieurs travaux de nature appliquée, touchant à des domaines variés, dans lesquels nous exploitons certaines des techniques mathématiques présentées précédemment, et particulièrement celles qui relèvent de la théorie de Perron-Frobenius non-linéaire et de la convexité tropicale. Ces applications utilisent aussi des techniques d'algèbre linéaire ou d'optimisation convexe.

English version

In this section, we describe several applied works in which we use some of the theoretical tools developed by the team, including non-linear Perron-Frobenius theory and tropical convexity. Some of these applications also make an intensive use of linear algebraic and convex programming methods.

6.6.2. Optimisation de la croissance de populations/Optimizing population growth

Participants: Vincent Calvez [ENS Lyon et Inria, NUMED], Pierre Gabriel [UVSQ], Stéphane Gaubert.

On s'intéresse dans [28] à l'optimisation du taux de croissance d'une population, représentée par un système dynamique $\dot{x}(t) = M(t)x(t)$, où la matrice $M(t)$ appartient à un ensemble compact de matrices de Metzler irréductibles. Ceci est motivé par un problème de biologie mathématique (modélisation de processus de croissance-fragmentation et protocole PMCA). Nous montrons que le taux de croissance est donné par la valeur propre non-linéaire d'un analogue max-plus de l'opérateur de Ruelle-Perron-Frobenius, ou de manière équivalente, par la constante ergodique d'une EDP d'Hamilton-Jacobi, dont les solutions et sous-solutions fournissent respectivement des normes de Barabanov et des normes extrémales. Nous exploitons les propriétés de contraction des flots monotones, relativement à la métrique projective de Hilbert, pour démontrer que le vecteur propre non-linéaire, qui correspond à une solution "KAM faible" de l'équation d'Hamilton-Jacobi, a bien une solution. Des exemples en petite dimension sont discutés, montrant en particulier que le contrôle optimal peut produire un cycle limite.

English version

We study in [28] a growth maximization problem for a continuous time positive linear system with switches. More precisely, we consider a dynamical system $\dot{x}(t) = M(t)x(t)$, where the matrix $M(t)$ must be chosen in a compact set of irreducible Metzler matrices. This is motivated by a problem of mathematical biology (modeling growth-fragmentation processes and the PMCA protocol). We show that the growth rate is determined by the non-linear eigenvalue of a max-plus analogue of the Ruelle-Perron-Frobenius operator, or equivalently, by the ergodic constant of a Hamilton-Jacobi (HJ) partial differential equation, the solutions or subsolutions of which yield Barabanov and extremal norms, respectively. We exploit contraction properties of order preserving flows, with respect to Hilbert's projective metric, to show that the non-linear eigenvector of the operator, or the "weak KAM" solution of the HJ equation, does exist. Low dimensional examples are presented, showing that the optimal control can lead to a limit cycle.

6.6.3. Preuve formelle d'inégalités non-linéaires/Formal proofs of non-linear inequalities

Participants: Xavier Allamigeon, Stéphane Gaubert, Victor Magron, Benjamin Werner [LIX].

La thèse de Victor Magron [153], dirigée par Benjamin Werner, codirigée par Stéphane Gaubert et Xavier Allamigeon, a porté sur la certification de bornes inférieures de fonctions multivariées à valeurs réelles, définies par des expressions semi-algébriques ou transcendantes, et sur la preuve de validité de celles-ci au moyen de certificats dans l'assistant de preuves Coq.

De nombreuses inégalités de cette nature apparaissent notamment dans la preuve par Thomas Hales de la conjecture de Kepler. Voici un exemple typique d'inégalité à prouver.

LEMME 9922699028 FLYSPECK. *Soit K, $\Delta\mathbf{x}$, l, t et f définis comme suit:*

$$\begin{aligned}
K &:= [4, 6.3504]^3 \times [6.3504, 8] \times [4, 6.3504]^2 \ , \\
\Delta\mathbf{x} &:= x_1x_4(-x_1 + x_2 + x_3 - x_4 + x_5 + x_6) \\
&\quad + x_2x_5(x_1 - x_2 + x_3 + x_4 - x_5 + x_6) \\
&\quad + x_3x_6(x_1 + x_2 - x_3 + x_4 + x_5 - x_6) \\
&\quad - x_2x_3x_4 - x_1x_3x_5 - x_1x_2x_6 - x_4x_5x_6 \ , \\
l(\mathbf{x}) &:= -\pi/2 + 1.6294 - 0.2213(\sqrt{x_2} + \sqrt{x_3} + \sqrt{x_5} + \sqrt{x_6} - 8.0) \\
&\quad + 0.913(\sqrt{x_4} - 2.52) + 0.728(\sqrt{x_1} - 2.0) \ , \\
t(\mathbf{x}) &:= \arctan \frac{\partial_4 \Delta\mathbf{x}}{\sqrt{4x_1\Delta\mathbf{x}}} \ , \\
f(\mathbf{x}) &:= l(\mathbf{x}) + t(\mathbf{x}) \ .
\end{aligned}$$

Alors, $\forall \mathbf{x} \in K, f(\mathbf{x}) \geq 0$.

On s'est donc intéressé à des fonctions non-linéaires, faisant intervenir des opérations semi-algébriques ainsi que des fonctions transcendantes univariées ($\cos$, $\arctan$, $\exp$, etc).

De manière classique, on peut approcher les fonctions transcendantes qui interviennent de la sorte par des polynômes, ce qui permet de se ramener à des problèmes d'optimisation semi-algébriques, que l'on peut résoudre par des techniques de sommes de carrés creuses conduisant à des problèmes SDP. Cependant, en pratique, cette approche est limitée par la taille des SDP à résoudre, qui croît rapidement avec le degré des approximations polynomiales.

Dans ce travail de thèse, on a développé une méthode alternative, qui consiste a borner certains des constituants de la fonction non-linéaire par des suprema de formes quadratiques dont les Hessiens sont judicieusement choisis. On reprend donc ici l'idée des approximations “max-plus” initialement introduites en contrôle optimal, en s'appuyant sur des techniques d'interprétation abstraite (généralisation non-linéaire de la méthode des gabarits de Manna et al.). Ainsi, on obtient une nouvelle technique d'optimisation globale, basée sur les gabarits, qui exploite à la fois la precision des sommes de carrés et la capacité de passage à l'échelle des méthodes d'abstraction.

L'implémentation de ces méthodes d'approximation a abouti à un outil logiciel : `NLCertify`. Cet outil génère des certificats à partir d'approximations semi-algébriques et de sommes de carrés. Son interface avec Coq permet de bénéficier de l'arithmétique certifiée disponible dans l'assistant de preuves, et ainsi d'obtenir des estimateurs et des bornes valides pour chaque approximation.

Les performances de cet outil de certification ont été démontrées sur divers problèmes d'optimisation globale ainsi que sur des inégalités essentiellement serrées qui interviennent dans la preuve de Hales (projet Flyspeck).

Ce travail est exposé dans [73], [74] et [18], [19].

English version

The PhD work of Victor Magron [153], supervised by Benjamin Werner, and cosupervised by Stéphane Gaubert and Xavier Allamigeon, dealt with the certification of lower bounds for multivariate functions, defined by semi-algebraic or transcendental expressions, and their correctness proof through certificates checked in the Coq proof assistant.

Many inequalities of this kind appear in particular in the proof by Thomas Hales of Kepler's conjecture. Here is a typical example of inequality.

LEMMA 9922699028 FLYSPECK. *Let K, $\Delta\mathbf{x}$, l, t and f be defined as follows:*

$$\begin{array}{rcl}
K & := & [4, 6.3504]^3 \times [6.3504, 8] \times [4, 6.3504]^2 \ , \\
\Delta\mathbf{x} & := & x_1 x_4(-x_1 + x_2 + x_3 - x_4 + x_5 + x_6) \\
 & & +x_2 x_5(x_1 - x_2 + x_3 + x_4 - x_5 + x_6) \\
 & & +x_3 x_6(x_1 + x_2 - x_3 + x_4 + x_5 - x_6) \\
 & & -x_2 x_3 x_4 - x_1 x_3 x_5 - x_1 x_2 x_6 - x_4 x_5 x_6 \ , \\
l(\mathbf{x}) & := & -\pi/2 + 1.6294 - 0.2213(\sqrt{x_2} + \sqrt{x_3} + \sqrt{x_5} + \sqrt{x_6} - 8.0) \\
 & & +0.913(\sqrt{x_4} - 2.52) + 0.728(\sqrt{x_1} - 2.0) \ , \\
t(\mathbf{x}) & := & \arctan \frac{\partial_4 \Delta\mathbf{x}}{\sqrt{4 x_1 \Delta\mathbf{x}}} \ , \\
f(\mathbf{x}) & := & l(\mathbf{x}) + t(\mathbf{x}) \ .
\end{array}$$

Then, $\forall \mathbf{x} \in K, f(\mathbf{x}) \geq 0$.

Thus, we considered non-linear functions, defined in terms of semi-algebraic operations and univariate transcendental functions ($\cos$, $\arctan$, $\exp$, etc).

Such transcendental functions can be classically approximated by polynomials, which leads to semi-algebraic optimization problems, which can be solved by sparse sum of squares techniques leading to SDP formulations. However, in practice, this approach is limited by the growth of the size of the SDP instances to be solved, whichs grows quickly with the degree of polynomial approximations.

In this PhD, we developed an alternative method, which consists in bounding some constituents of the non-linear function to be optimized by suprema of quadratic forms with well chosen Hessians. This is based on the idea of "maxplus approximation" initially introduced in optimal control, and also, on abstract interpretation (the template method introduced by Manna et al. in static analysis). In this way, we end up with a new global optimization technique, which takes advantage of the precision of sum of squares and of the scalability of abstraction methods.

These methods have been implemented in a software tool: `NLCertify`. This tool generates certificates from semi-algebraic and sum of square certificates. Its interface with Coq allows one to take benefit of the certified arithmetics available in this proof assistant, and so, to obtain estimators and valid bounds for each approximation.

The performances of this certification tool have been shown on several global optimization problems from the literature, as well as on essentially tight inequalities taken from Hales' proof (Flyspeck project).

This work is presented in [73], [74] and [18], [19].

6.6.4. Vérification de systèmes temps-réels/Verification of real-time systems

Participants: Xavier Allamigeon, Uli Fahrenberg [IRISA], Stéphane Gaubert, Ricardo Katz [Conicet], Axel Legay [IRISA].

Dans [152], Lu, Madsen, Milata, Ravn, Fahrenberg et Larsen ont montré que les polyèdres tropicaux peuvent être utilisés dans le cadre de l'analyse d'accessibilité d'automates temporisés. En effet, les polyèdres tropicaux expriment naturellement des invariants non-convexes, qui sont en fait des disjonctions d'invariants fournis par des DBM (*difference bound matrices*). A ce titre, les polyèdres tropicaux devraient permettre de réduire le nombre de disjonctions réalisées pendant l'analyse d'automates temporisés. Une limitation importante de cette approche est cependant que les polyèdres tropicaux sont topologiquement fermés, et qu'ils ne peuvent donc pas exprimer de contraintes d'inégalités strictes. Ces dernières sont néanmoins fondamentales dans l'analyse de systèmes temps-réels.

Nous avons donc développé dans [69] une généralisation des polyèdres tropicaux permettant d'exprimer des contraintes mixtes, *i.e.* strictes ou larges. Notre approche repose sur l'utilisation d'inégalités tropicales linéaires à coefficients dans un (quotient du) semi-anneau de germes affines. Afin de réaliser des opérations sur cette nouvelle classe de polyèdres tropicaux, nous avons défini deux nouveaux algorithmes. Le premier est un analogue tropical de l'élimination de Fourier-Motzkin. Celle-ci s'applique plus généralement à des systèmes d'inégalités linéaires sur des semi-anneaux idempotents et totalement ordonnés. Le second algorithme permet de tester si un système de contraintes mixtes admet une solution. Nous montrons en effet que ce problème est équivalent en temps polynomial à la résolution d'un problème de jeux déterministes à somme nulle. Ces deux contributions nous permettent de définir les primitives requises pour l'analyse d'accessibilité d'automates temporisés.

English version

Lu, Madsen, Milata, Ravn, Fahrenberg and Larsen have shown in [152] that tropical polyhedra can be applied to the reachability analysis of timed automata. Indeed, tropical polyhedra naturally express non-convex invariants, which correspond to disjunctions of invariants provided by DBM (*difference bound matrices*). Consequently, tropical polyhedra should allow to reduce the number of disjunctions arising during the analysis of timed automata. An important limitation of this approach is that tropical polyhedra are topologically closed, and thus they cannot express strict inequality constraints. However, such constraints plays an important role in the analysis of real-time systems.

As a result, we have developed in [69] a generalization of tropical polyhedra, in order to express mixed constraints, *i.e.* strict or loose ones. Our approach relies on tropical linear inequalities with coefficients in a (quotient of) the semiring of affine germs. In order to perform operations on this new class of polyhedra, we have introduced two new algorithms. The first one is a tropical analog of Fourier-Moztkin elimination. In fact, it applies more generally to systems of linear inequalities over totally ordered and idempotent semirings. The second algorithm allows to test the feasability of a mixed constraint system. We indeed show that this problem is polynomial-time equivalent to solving mean payoff games. These two contributions allow to define the primitives required by the reachability analysis of timed automata.

6.6.5. Géométrie de l'ordre de Loewner et application au calcul d'invariants quadratiques en analyse statique de programme/Geometry of the Loewner order and appliction to the synthesis of quadratic invariants in static analysis of program

Participants: Xavier Allamigeon, Stéphane Gaubert, Éric Goubault [LIX], Sylvie Putot [LIX], Nikolas Stott.

Le stage de recherche de l'École des Mines de Nikolas Stott a porté sur la caractérisation de l'ensemble des majorants minimaux de deux matrices symmétriques, relativement à l'ordre de Loewner, et sur l'application de cette caractérisation à la synthèse d'invariants quadratiques en analyse statique de programme.

English version

The research internship of “École des Mines” made by Nikolas Stott dealt with the characterization of the set of minimal upper bounds of two matrices with respect to Loewner order, motivated by the generation of quadratic invariants in static analysis of programs.

6.6.6. Optimisation de l’affectation temps réel des moyens de secours des pompiers/Optimization of the real time assignment of firemen vehicles

Participants: Marianne Akian, Xavier Allamigeon, Vianney Boeuf, Stéphane Gaubert, Stéphane Raclot [BSPP].

La thèse de Vianney Boeuf, qui a démarré en Septembre, est effectuée en partenariat avec la Brigade des Sapeurs Pompiers de Paris (BSPP). Elle est motivée par l’optimisation des moyens de secours, en incluant les questions de dimensionnement et d’affectation temps réel des moyens. On s’intéresse en particulier à l’affectation des engins et véhicules de secours, éventuellement empruntés à différentes casernes. Ce travail intervient en complément du travail de l’équipe au sein du projet ANR Democrite, qui porte sur l’évaluation du risque en milieu urbain.

English version

The PhD work of Vianney Boeuf started in September. It is carried out with the Brigade of Paris Firemen (BSPP). It is motivated by the issue of optimization of emergency resources, including the real time dynamic assignment of engines or emergency vehicles. This work is carried out in complement to the ANR project Democrite, dealing with risk evaluation in urban environment.

MC2 Team

6. New Results

6.1. Highlights of the Year

- **Models for gliomas**
  Glioblastoma multiforme (GBM) causes significant neurological morbidity and short survival times. Brain invasion by GBM is associated with poor prognosis. Recent clinical trials of bevacizumab in newly-diagnosed GBM found no beneficial effects on overall survival times; however, the baseline health-related quality of life and performance status were maintained longer in the bevacizumab group and the glucocorticoid requirement was lower. In a recent work in collaboration with UAB, we have constructed a clinical-scale model of GBM whose predictions uncover a new pattern of recurrence in 11/70 bevacizumab-treated patients. The findings support an exception to the Folkman hypothesis: GBM grows in the absence of angiogenesis by a cycle of proliferation and brain invasion that expands necrosis. Furthermore, necrosis is positively correlated with brain invasion in 26 newly-diagnosed GBM. The unintuitive results explain the unusual clinical effects of bevacizumab and suggest new hypotheses on the dynamic clinical effects of migration by active transport, a mechanism of hypoxia-driven brain invasion.
- **Electroporation modeling** (M. Leguebe, C. Poignard)
  Based on the new discovery of the team of Vectorolgy and anti-cancerous therapies on the membrane lipid oxidation during the pulse delivery, we have provided a model of cell permeabilization that makes it possible to explain the process of electroporation : pore formation during the pulse and surface diffusion of altered lipids after the pulse. Our model explains the long-term effect of electroporation (the permeable state of the membrane lasts a few minutes after the pulse delivery). A 3D-code in C++ has been implemented during the PhD thesis of M. Leguèbe. The team MC2 is now part of the European Lab EBAM on electroporation modeling. An international workshop on Electroporation and Biophysical Therapies was held in Bordeaux the 15th and 16th December.
- Simulation of **multiphysic fluid-structure impacts in 3D**. See http://www.math.u-bordeaux1.fr/ adebrauer/ for astinishing videos.

6.2. Cancer modeling

- **Patient specific simulation for lung metastases**
  The calibration process has tremendously improved by a deep study of the model and its parameter space. Work is ongoing to validate the whole process on a retrospective study of 30 patients. A prototype is being built for our collaborators at Institut Bergonié to use in their clinical routine. The same strategy has been applied to meningiomas in the last year of the post-doc of Julie Joie within the IRL MONICA with a retrospectuve study on 10 patients.
- Modelling of the response to targeted therapies for liver metastasis of a gist : 2 clinical cases with a long term longitudinal follow-up with CT-scans. We are able to fit the volume of the lesion but also the the texture of the image, that is the ratio between necrotic tissues and proliferative ones. See [82].
- Tumor growth model for ductal carcinoma: from in situ phase to stroma invasion. See [71].
- Permeable and conducting states of membrane submitted to electric pulse: non-linear PDE model, 2D and 3D code in C++.
- Free boundary value model for invadopodia and migration of cell developed in collaboration with Osaka University and Tokyo University of Sciences.
- Endothelial cell migration on polymers: agent based model. Paper accepted in DCDS-B.

- A. Peretti started her PhD on the modeling of the heterogenity on renal cancer.
- Benjamin Taton started a post-doc on the modeling of the renal function through perfusion MRI. B. Taton is a MD.
- Th. Michel obtained some mathematical properties on the system of PDEs used for the modelling of GIST metastases.
- **Models for preclinical studies**
  - Mathematical **ODE models of tumor volume kinetics in mice** (collaboration with the Center of Cancer and Systems Biology, Boston, USA and J. Ebos, Roswell Park Cancer Institute, Buffalo, USA).
    Rational and quantitative evaluation of the predictive and descriptive power of the majority of the classical ODE models for tumor growth against data from two distinct experimental systems [57]. One of the major finding was the huge improvement of the predictive properties when using the population *a priori* information on the distribution of the parameters.
  - Mathematical model for data of **preclinical metastatic burden dynamics** and clinical data of metastatic relapse probability of **breast cancer** (collaboration with J. Ebos, Roswell Park Cancer Institute, Buffalo, USA).
    Validation of the descriptive and predictive ability of a simple and minimally parameterized model. The major finding resulting from the modeling analysis was the quantification of the impact of surgery on survival improvement (highly nonlinear), which suggests a threshold primary tumor size for efficacy of the surgery in terms of preventing metastatic recurrence. A publication is in preparation.
  - Effect of **anti-cancer therapies** in preclinical experiments
    * Evaluation of several models (several already published but also new ones) for the effect of **anti-angiogenic drugs**[0] on tumor growth, based on statistical parameter estimation methods on experimental data (collaboration with J. Ebos, Roswell Park Cancer Institute, Buffalo, USA). The main finding was one model that was able to both describe the effect of the drug (Sunitinib) and predict the effect when changing the scheduling. See [66].
    * Effect of the **sequence of administration between cytotoxic and anti-angiogenic drugs** (collaboration with J. Ciccolini and D. Barbolosi, SMARTc, Inserm, Marseille, Fr). See [84].
- **Theoretical cancer biology**
  - Theories of metastatic initiation (collaboration with A. Bikfalvi, LAMC, Inserm and the RMSB, CNRS in Bordeaux, Fr).
    Confrontation of theories and experimental data challenged the classical view of metastatic establishment and growth and suggested that tumors could merge in initial phases. Quantitative impact of the merging was studied using a dedicated and properly calibrated spatial model.
  - Tumor-tumor distant interactions (collaboration with the Center of Cancer and Systems Biology, Boston, USA).
    Statistical and modeling analysis of experimental data for two tumors implanted in one organism.

[0] recent anti-cancer drugs that target the tumor vasculature rather than the cancer cells themselves

6.3. Newtonian fluid flows simulations and their analysis

- Development of a high-order (third order in time and space) level-set method which allow to compute consistently the curvature of the interface even for long times (L. Weynans, F. Luddens and M. Bergmann)
- Development of a sharp cartesian method for the simulation of incompressible flows with high density ratios, like air-water interfaces. This method is inspired from the second-order cartesian method for elliptic problems with immersed interfaces developed in Cisternino-Weynans [69]
- Study of the convergence in 1D and 2D of the method developed in Cisternino-Weynans [69]

MCTAO Project-Team

6. New Results

6.1. Optimal control for quantum systems and NMR

Participants: Bernard Bonnard, Mathieu Claeys [Imperial College, UK], Olivier Cots, Thierry Combot, Pierre Martinon [project team COMMANDS], Alain Jacquemard [Université de Bourgogne, IMB].

- The contrast imaging problem in nuclear magnetic resonance can be modeled as a Mayer problem, in the terminology of optimal control. The candidates as minimizers are selected among a set of extremals, solutions of a Hamiltonian system given by the Pontryagin Maximum Principle; sufficient second order conditions are known; they form the geometric foundations of the HAMPATH code which combines shooting and continuation methods.

  In [4], based on these theoretical studies, a thorough analysis of the case of deoxygenated/oxygenated blood samples is pursued, based on many numerical experiments.
- We initiated more than a year ago a program to compare and study the complementarities between these methods based on the Pontryagin Maximum Principle are known as indirect methods,
  - with the so-called direct methods where optimal control is seem as a generic optimization problem, as implemented in the Bocop software, developed in the COMMANDS project-team,
  - and with LMI techniques used to obtain global bounds on the extremum;
  this was naturally done in collaboration with Pierre Martinon, an important contributor to Bocop and with Mathieu Claeys (LAAS CNRS, a PhD student supervised by J.-B. Lasserre, now with Imperial College). The results are very promising, and there is a gain, numerically, in using both direct and indirect methods while working towards global optimality (in the contrast problem there are many local optima and the global optimality is a complicated issue). This is presented in [3].

  This also led to use algebraic techniques to further analyse the equations and their dependance of the materials to be discriminated [10].
- For time minimal control of a linear spin system with Ising coupling (more complex than the model above), we also analysed *integrability* properties of extremal solutions of the Pontryagin Maximum Principle, in relation with conjugate and cut loci computations. Restricting to the case of three spins, as in [11], the problem is equivalent to analyze a family of almost-Riemannian metrics on the sphere S^2 , with Grushin equatorial singularity. The problem can be lifted into a SR-invariant problem on $SO(3)$, this leads to a complete understanding of the geometry of the problem and to an explicit parametrization of the extremals using an appropriate chart as well as elliptic functions. This approach is compared with the direct analysis of the Liouville metrics on the sphere where the parametrization of the extremals is obtained by computing a Liouville normal form. This is backed by an algebraic approach applying differential Galois theory to integrability.

6.2. Conjugate and cut loci computations and applications

Participants: Bernard Bonnard, Olivier Cots, Jean-Baptiste Caillau, Alessio Figalli [Univ. of Texas at Austin, USA], Thomas Gallouët [MEPHYSTO project-team], Ludovic Rifford.

- Many optimal control problems from mechanics or quantum systems (see [11] and the last paragraph of section 6.1) lead to studying some king of singular metrics, sometimes known as almost-Riemannian. This led us to consider, in [2], metrics on the two-sphere of revolution of the following find: they are Riemannian on each open hemisphere whereas one term of the corresponding tensor becomes infinite on the equator. Length minimizing curves can be computed and structure results on the cut and conjugate loci given, extending those in [25]. These results rely on monotonicity and convexity properties of the quasi-period of the geodesics; such properties are studied on an example with elliptic transcendency. A suitable deformation of the round sphere allows to reinterpretate the equatorial singularity in terms of concentration of curvature and collapsing of the sphere.

- It is known that convexity of the injectivity domain (the boundary of which is sent by the exponential map to the first cut locus) and the "Ma–Trudinger–Wang condition" (an positivity condition on the Ma–Trudinger–Wang tensor) both play a very important role in the continuity of solutions of optimal transport problems. This led to study these properties on their own, and it is still an open question to decide under which conditions the latter implies the former. In [13], it is proved that the MTW condition implies the convexity of injectivity domains on a smooth nonfocal compact Riemannian manifold. This improves a previous result by Loeper and Villani.

6.3. Averaging in control and application to space mechanics

Participants: Bernard Bonnard, Helen-Clare Henninger, Jana Němcová [Institute of Chemical Tech, Prague, CZ], Jean-Baptiste Pomet, Jeremy Rouot.

As explained in sections 3.5 and 4.1 , control problems where the non controlled system is conservative and the control effect is small compared to the free dynamics lead to computing an average system. This computation may be explicit or numerical.

Even though it will not be always the case that an explicit expression is available, it is interesting to study that case thoroughly.

- In [23], [24], a smooth Riemannian metric was introduced to describe the energy minimizing orbital transfer with low propulsion. We have pursued a study of its deformation due to the standard perturbations in space mechanics, e.g. oblate spheroid shape of the Earth and lunar attraction. In [12], using Hamiltonian formalism, we describe the effects of the perturbations on the orbital transfers and the deformation of the conjugate and cut loci of the original metric. This is done using averaging with respect to both the proper frequency of the space vehicle and the moon frequency.
- The average system has the advantage of being more controllable (it has new virtual controls), but often displays singularities that were not present in the original system. It is the case when minimum time is considered instead of the quadratic energy criterium. We are conducted an analysis of this average minimum time Hamiltonian flow.

  In [6], we compare the two problems for planar transfers. While the energy case leads to analyze a 2D Riemannian metric using the standard tools of Riemannian geometry (curvature computations, geodesic convexity), the time minimal case is associated to a Finsler metric which is not smooth. Nevertheless a qualitative analysis of the geodesic flow is given in this article to describe the optimal transfers. In particular we prove geodesic convexity of the elliptic domain.

6.4. Applications of control methods to dynamical systems

Participants: Gonzalo Contreras, Alessio Figalli, Ayadi Lazrag, Ludovic Rifford, Raffael Ruggiero.

Ludovic Rifford and collaborators have been applying with success, techniques from geometric control theory to open problems in dynamical systems, mostly on genericity properties and using controllability methods to build suitable perturbations.

This has been applied to closing geodesics and weak-KAM theory [39], [38].

Ayadi Lazrag's PhD also deals with such problems; applying techniques close to these in [61], he established a version of Francks' lemma for geodesic flows; one goal is to apply this to persistence problems. The approach relies on control theory results, with order 2 conditions. See [14] and [15], where a non trivial conjecture on generic hyperbolicity of the so-called Aubry set of a Hamiltonian is solved on compact surfaces and in the C^2 topology (for genericity).

MEPHYSTO Team

6. New Results

6.1. Highlights of the Year

A. Gloria, S. Neukamm, and F. Otto published their recent contribution [17] on quantitative homogenization in Inventiones Mathematicae.

As a plenary speaker of the World Congress of Computational Mechanics in Barcelone in July 2014, P. Le Tallec (Ecole polytechnique) presented our joint results [15], [25].

6.2. Quantitative stochastic homogenization

A. Gloria, S. Neukamm (Univ. Dresden), and F. Otto (MPI for mathematics in the sciences, Leipzig) developed in [17] a general approach to quantify ergodicity in stochastic homogenization of discrete elliptic equations. Using a parabolic approach, they obtained optimal estimates on the time-decay of the so-called environment seen from the particle. This allowed them to prove optimal bounds on the corrector gradient and the corrector itself in any dimension (thus improving on [4]). They also obtained the first error analysis of the popular periodization method to approximate the homogenized coefficients.

In [32], A. Gloria and F. Otto extended their results [4], [5] on discrete elliptic equations to the continuum setting. They treated in addition the case of non-symmetric coefficients, and obtained optimal estimates in all dimensions by the elliptic approach (whereas [4], [5] were suboptimal for $d = 2$).

In [28], A. Gloria and D. Marahrens (MPI for mathematics in the sciences, Leipzig) extended the annealed results [51] on the discrete Green function by D. Marahrens and F. Otto to the continuum setting. As a by-product of their result, they obtained new results in uncertainty quantification by estimating optimally the variance of the solution of an elliptic PDE whose coefficients are perturbed by some noise with short range of dependence.

In their recent work [29], A. Gloria, S. Neukamm, and F. Otto developed a regularity theory for random elliptic operators inspired by the contributions of Avellaneda and Lin [39] in the periodic setting and of our visitor S. Armstrong with C. Smart [38]. This allowed them to consider coefficients with arbitrarily slow decaying correlations in the form of a family of correlated Gaussian fields.

In [30], A. Gloria and J. Nolen (Duke Univ.) proved a quantitative central limit theorem for the effective conductance on the discrete torus. In particular, they quantified the Wasserstein distance between a normal random variable and the CLT-like rescaling of the difference between the approximation of the effective conductance by periodization and the effective conductance. Their estimate is sharp and shows that the Wasserstein distance goes to zero (up to logarithmic factors) as if the energy density of the corrector was iid (which it is not). This completes and settles the analysis started in [17] on the approximation of homogenized coefficients by periodization by characterizing the limiting law in addition to the scaling.

6.3. Derivation of nonlinear elasticity from polymer-physics

In [15], A. Gloria, P. Le Tallec (Mechanics department, Ecole polytechnique), and M. Vidrascu (Project-team REO, Inria) numerically investigated the nonlinear elasticity model obtained in [1] by discrete stochastic homogenization, and compared it to the standard measurements by Treloar on natural rubber. The implementation was realized in the Modulef software. These results are in rather good agreement, which shows that the approach seems to be promising.

In [25], M. de Buhan (CNRS, Univ. Paris Descartes), A. Gloria, P. Le Tallec and M. Vidrascu proposed a numerical method to produce analytical approximations (that can be used in practical nonlinear elasticity softwares) of the numerical approximations obtained in [15] of the discrete-to-continuum energy density derived theoretically in [1]. This numerical method is based on the parametrization of the set of polyconvex Ogden laws and on the combination of a least square method and a genetic algorithm (cf. CMA-ES).

6.4. Numerical homogenization

Inspired by the quantitative analysis of [17] and [48], Z. Habibi (former SIMPAF post-doctoral fellow) and A. Gloria introduced in [14] a general method to reduce the so-called resonance error in numerical homogenization, both at the levels of the approximation of the homogenized coefficients and of the correctors. This method significantly extends [2]. The method relies on the introduction of a massive term in the corrector equation and of a systematic use of Richardson extrapolation. In the three academic examples of heterogeneous coefficients (periodic, quasiperiodic, and Poisson random inclusions), the method yields optimal theoretical and empirical convergence rates, and outperforms most of the other existing methods.

6.5. Nonlinear Schrödinger equation

S. De Bièvre, S. Rota Nodari (CEMPI postdoc 2013-2015) and F. Genoud (CEMPI visitor, September 2013) have explained the geometry underlying the so-called energy-momentum method for proving orbital stability in infinite dimensional Hamiltonian systems. Applications include the orbital stability of solitons of the NLS and Manakov equations. This work is to appear as a chapter (120p) in the first volume of the CEMPI Lecture Notes in Mathematics, cf. [24].

6.6. Kicked rotors

S. De Bièvre and his PhD student E. Soret rigorously proved the growth rate of the energy in a Markovian model for stochastic acceleration of a particle in a random medium, cf. [34].

6.7. Time integration of Hamiltonian system with noise

G. Dujardin introduced an energy preserving method for Hamiltonian dynamics perturbed by a multiplicative noise, cf. [11].

6.8. Miscellaneous results

The MEPHYSTO team is currently hosting former members of the project-team SIMPAF who focus on numerical methods for dissipative systems:

- corrosion models [19], [23],
- fluid mechanics [9], [21], [27], [10],
- numerical analysis for asymptotic preserving properties [8], [7],
- a posteriori estimates [20].

T. Gallouët also made contributions in optimal transport [22], [26].

MESCAL Project-Team

6. New Results

6.1. Simulation of distributed architectures

- Simgrid is a toolkit providing core functionalities for the simulation of distributed applications in heterogeneous distributed environments. It models fine-grain detail of the studied platform. In [6], we present quantitative results that show that SimGrid compares favorably to state-of-the-art domain-specific simulators in terms of scalability, accuracy, or the trade-off between the two. In [37], [29], we develop an hybrid approach of simulation and emulation of applications that use starPU. By using this approach, Simgrid calibrates the time to run specific subtasks at runtime and simulates all system calls of the application. This approach allows us to obtain performance results that are within one percent of measured results.
- In [33], [18], we study the problem of sampling the stationary distribution of a random walker in $\{0 \cdots N\}^d$ using simulation. This algorithm combines the rejection method and coupling from the past of a set of trajectories of the Markov chain that generalizes the classical sandwich approach. We also provide a complexity analysis of this approach in several cases showing a coupling time in $O(N^2 d \log d)$ when no arc is forbidden and an experimental study of its performance.

6.2. Interactive Analysis and Visualization of Large Distributed Systems

- In [13], we review the methodology that we use to visualize information for large-scale data-set. Our approach uses tools from information theory to define a trade-off between the loss of information and the compactness of the representation. This methodology is applied to spatio-temporal representation of traces of execution in [30], [16], [17], [32]. In these papers, we show how to build a concise overview of the trace behavior as the result of a spatio-temporal data aggregation process. The experimental results show that this approach can help the quick and accurate detection of anomalies in traces containing up to two hundred million events.
- Trace analysis graphical user environments have to provide different views on trace data, to really help provide insights on the traced application behavior. In [22], [35], we propose an open and modular software architecture, the FrameSoC workbench, that defines clear principles for view engineering and for view consistency management. The FrameSoC workbench has been successfully applied in real trace analysis use-cases. This work has also been tested on real scenario coming from a collaboration with ST Microelectronic [25].
- In [7], we design a novel prediction method with Bayes model to predict a load fluctuation pattern over a long-term interval, in the context of Google data centers. All of the prediction methods are evaluated using Google trace with 10,000+ heterogeneous hosts. Experiments show that our Bayes method improves the long-term load prediction accuracy by up 5 to 50%, compared to other state-of-the-art methods.

6.3. Management of Parallel Architectures

- In [12], we present a topology-aware load balancing algorithm for parallel multi-core machines and its proof of asymptotic convergence to an optimal solution. The algorithm, named HwTopoLB, takes into account the properties of current parallel systems composed of multi-core compute nodes, namely their network interconnection, and their complex and hierarchical core topology. We have implemented HwTopoLB using the Charm++ Parallel Runtime System and evaluated its performance with two different benchmarks and one application. Our experimental results confirms that HwTopoLB outperform existing load balancing strategies on different multi-core systems.

- Large scale distributed systems typically comprise hundreds to millions of entities that have only a partial view of resources. How to fairly and efficiently share such resources between entities in a distributed way has thus become a critical question. In [31], we develop a possible answer based on Lagrangian optimization and distributed gradient descent. Under certain conditions, the resource sharing problem can be formulated as a global optimization problem, which can be solved by a distributed self-stabilizing demand and response algorithm.
- The management of resources on testbeds, including their description, reservation and verification, is a challenging issue, especially on of large scale testbeds such as those used for research on High Performance Computing or Clouds. In [23], we present the solution designed for the Grid'5000 testbed in order to: (1) provide users with an in-depth and machine-parsable description of the testbed's resources; (2) enable multi-criteria selection and reservation of resources using a HPC resource manager; (3) ensure that the description of the resources remains accurate. In [24], we present Kascade, a solution for the broadcast of data to a large set of compute nodes. We evaluate Kascade using a set of large scale experiments in a variety of experimental settings, and show that Kascade: (1) achieves very high scalability by organizing nodes in a pipeline; (2) can almost saturate a 1 Gbit/s network, even at large scale; (3) handles failures of nodes during the transfer seamlessly because of its fault-tolerant design.

6.4. Reproducible experiments and papers

- In the field of large-scale distributed systems, experimentation is particularly difficult. The studied systems are complex, often nondeterministic and unreliable, software is plagued with bugs, whereas the experiment workflows are unclear and hard to reproduce. In [5], we provide an extensive list of features offered by general-purpose experiment management tools dedicated to distributed systems research on real platforms. We then use it to assess existing solutions and compare them, outlining possible future paths for improvements.
- Experiment reproducibility is a milestone of the scientific method. Reproducibility of experiments in computer science would bring several advantages such as code re-usability and technology transfer. The reproducibility problem in computer science has been solved partially, addressing particular class of applications or single machine setups. In [26], we present our approach oriented to setup complex environments for experimentation, environments that require a lot of configuration and the installation of several software packages. The main objective of our approach is to enable the exact and independent reconstruction of a given software environment and the reuse of code. We present a simple and small software appliance generator that helps an experimenter to construct a specific software stack that can be deployed on different available testbeds. [14],
- In [28], [45], we address the question of developing a lightweight and effective workflow for conducting experimental research on modern parallel computer systems in a reproducible way. Our workflow simply builds on two well-known tools (Org-mode and Git) and enables us to address issues such as provenance tracking, experimental setup reconstruction, replicable analysis. Although this workflow is perfectible and cannot be seen as a final solution, we have been using git for two years now and we have recently published a fully reproducible article, which demonstrates the effectiveness of our proposal.

6.5. Game Theory and Distributed Optimization

- In wireless networks, channel conditions of and user quality of service (QoS) requirements vary, often quite arbitrarily, with time (e.g. due to user mobility, fading, etc.) In this dynamic setting, static solution concepts (such as Nash equilibrium) are no longer relevant. Hence, we focus on the concept of no-regret : policies that perform at least as well as the best fixed transmit profile in hindsight. In [21], we examine the performance of the seminal Foschini–Miljanic (FM) power control scheme in a random environment. We provide a formulation of power control as an online optimization problem and we show that the FM dynamics lead to no regret in this dynamic context. We introduce an adjusted version of the FM algorithm which retains the convergence and no-regret properties of the original algorithm in this constrained setting. In [27], we examine the problem of cost / energy-efficient power allocation in uplink multi-carrier orthogonal frequency-division multiple access wireless networks. We use tools from stochastic convex programming to develop a learning scheme that retains its convergence properties irrespective of the magnitude of the observational errors. In [11], we consider a cognitive radio network where wireless users with multiple antennas communicate over several non-interfering frequency bands. We draw on the method of matrix exponential learning and online mirror descent techniques to derive a no-regret policy that relies only on local channel state information.

- In game theory, the best-response strategy of a player is a strategy that maximizes the selfish payoff of this player. A natural and popular question is, when players update their strategy over time, do they converge to a Nash equilibrium. In [15], we characterize the revision sets in different variants of the best response algorithm that guarantee convergence to pure Nash Equilibria in potential games. We prove that if the revision protocol is separable, then the greedy version as well as smoothed versions of the algorithm converge to pure Nash equilibria. If the revision protocol is not separable, then convergence to Nash Equilibria may fail in both cases. In [43], we investigate a class of reinforcement learning dynamics in which each player plays a "regularized best response" to a score vector consisting of his actions' cumulative payoffs. Our main results extend several properties of the replicator dynamics such as the elimination of dominated strategies, the asymptotic stability of strict Nash equilibria and the convergence of time-averaged trajectories to interior Nash equilibria in zero-sum games.

6.6. Agent-based modeling and applications to Smart Energy and Transportation Systems

- Renewable energy sources, such as wind, are characterized by non-dispatchability, high volatility, and non-perfect forecasts. Energy storage or electric loads that have a flexible consumption are viewed as a way to mitigate these effects. In [9], [19], we study centralized and distributed algorithms for solving this problem. We provide theoretical bounds on the trade-off between energy loss and the use of reserves. We develop a centralized algorithm that attains this bound in [9]. In [19], we study a distributed optimization problem by modeling a two-stage electricity market. We show that the market is efficient: the players' selfish responses to prices coincide with a socially optimal policy. We develop a distributed solution technique based on the Alternating Direction Method of Multipliers (ADMM) and trajectorial forecasts to compute the Nash-equilibrium.
- Bike-sharing systems are becoming important for urban transportation. In these systems, users arrive at a station, pick up a bike, use it for a while, and then return it to another station of their choice. In [8], we propose a stochastic model of an homogeneous bike-sharing system and study the effect of the randomness of user choices on the number of problematic stations. Even in a homogeneous city, the system exhibits a poor performance: the minimal proportion of problematic stations is of the order of the inverse of the capacity. We show that simple incentives, such as suggesting users to return to the least loaded station among two stations, improve the situation by an exponential factor.
- In [10], we discuss the validation of an agent-based model of emergent city systems with heterogeneous agents. We transform our model into an analytically tractable discrete Markov model, and we examine the city size distribution. We show that the Markov chains lead to a power-law distribution when the ranges of migration options are randomly distributed across the agent population. We also identify sufficient conditions under which the Markov chains produce the Zipf's Law, which has never been done within a discrete framework. The conditions under which our simplified model yields the Zipf's Law are in agreement with, and thus validate, the configurations of the original heterogeneous agent-based model.

MEXICO Project-Team

6. New Results

6.1. Highlights of the Year

6.1.1. Active Diagnosis for Probabilistic Systems

Diagnosis fits well with probabilistic systems since it is natural to model the uncertainty about the behaviour of a partially observed system by distributions. We had previously revisited the active diagnosis (which aims at controlling the system to make it diagnosable) in discrete event systems designing optimal decision and synthesis procedures [7]. This year, we have considered active diagnosis for probabilistic discrete event systems, obtaining again optimal procedures [26]. Furthermore we have refined the notion of active diagnosis by introducing the *safe active diagnosis* which ensures that after the control is applied, there is a positive probability that a fault never occurs. Interestingly this problem is undecidable but for finite memory controller we have shown that the problem becomes again decidable and we have designed optimal decision and synthesis procedures. Our approach has raised an issue that has not be observed by previous researchers: while in discrete event system, most variants of diagnosis are in fact equivalent, this is no more the case for probabilistic systems. So in [26], we have undertaken the task of classifying the different versions obtaining a complete landscape of the notions both in terms of relations and complexity. Furthermore we have proposed a new notion of diagnosis, the *prediagnosis* that combines the advantages of diagnosis and prediction.

6.1.2. Weighted automata and weighted logics

Weighted automata are a conservative quantitative extension of finite automata that enjoys applications, e.g., in language processing and speech recognition. Their expressive power, however, appears to be limited, especially when they are applied to more general structures than words, such as graphs. To address this drawback, we have introduced weighted pebble walking automata, which allow to navigate freely in the graph and may use pebbles to mark some positions.

In [20], we have shown with examples from natural language modeling and quantitative model-checking that weighted expressions and automata with pebbles are more expressive and allow much more natural and intuitive specifications than classical ones. We have extended Kleene-Schu ̈tzenberger theorem showing that weighted expressions and automata with pebbles have the same expressive power. We focussed on an efficient translation from expressions to automata. We also proved that the evaluation problem for weighted automata can be done very efficiently if the number of reusable pebbles is low.

In [18], we have studied the expressive power of these automata on words. We have proved that two-way pebble weighted automata, one-way pebble weighted automata, and our weighted logic with transitive closure are expressively equivalent. We also gave new logical characterizations of standard recognizable series.

In [30], we addressed the more general case of graphs such as nested words, trees, pictures, Mazurkiewicz traces, ... We established that weighted pebble walking automata have the same expressive power as weighted first order logic with transitive closure logic, lifting a similar result by Engelfriet and Hoogeboom from the Boolean case to a quantitative setting.

6.1.3. Verification of concurrent recursive programs

Distributed systems form a crucially important but particularly challenging domain. Designing correct distributed systems is demanding, and verifying its correctness is even more so. The main cause of difficulty here is concurrency and interaction (or communication) between various distributed components. Hence it is important to provide a framework that makes easy the design of systems as well as their analysis. There are two schools of thought on reasoning about distributed systems: one following the interleaving based semantics, and one following the visual partial-order/graph based semantics. In [23], we compare these two approaches and argue in favour of the latter. An introductory treatment of the split-width technique is also provided.

In [34], we develop a general technique based on split-width for the verification of networks of multi-threaded recursive programs communicating via reliable FIFO channels. We extend the approach of [6] to this setting. Split-width offers an intuitive visual technique to decompose our behaviour graphs such as MSCs and nested words. The decomposition is mainly a divide-and-conquer technique which naturally results in a tree decomposition. Every behaviour can now be interpreted over its decomposition tree. Properties over the behaviour naturally transfer into properties over the decomposition tree. This allows us to use tree-automata techniques to obtain decision procedures for a range of problems such as reachability, model checking against logical formalisms etc. In this way, we obtain simple, uniform and optimal decision procedures for various verification problems parametrised by split-width. Furthermore, the simple visual mechanism of split-width is as powerful as yardstick graph measures such as tree-width or clique-width. Hence it captures any class of distributed behaviours with a decidable MSO theory.

Multi-threaded recursive programs communicating via channels are turing powerful, hence their verification has focussed on under-approximation techniques. Any error detected in the under-approximation implies an error in the system. However the successful verification of the under-approximation is not as useful if the system exhibits unverified behaviours. In [24], we study controllers that observe/restrict the system so that it stays within the verified under-approximation. We identify some important properties that a good controller should satisfy. We consider an extensive under-approximation class, construct a distributed controller with the desired properties and also establish the decidability of verification problems for this class.

6.1.4. Regulation in Systems Biology

6.1.4.1. Rare events in Signalling Cascades

The visit in 2013 of Professor Monika Heiner from Cottbus University has led to a fruitful collaboration related to statistical model checking of rare events in signalling cascades (a regulatory biological system) [25]. This work has received one of the five top paper awards of the conference. In addition, we have improved the statistical methods used in our tool Cosmos.

6.1.4.2. Characterization of Reachable Attractors Using Petri Net Unfoldings

Attractors of network dynamics represent the long-term behaviours of the modelled system. Their characterization is therefore crucial for understanding the response and differentiation capabilities of a dynamical biological system. In the scope of qualitative models of interaction networks, the computation of attractors reachable from a given state of the network faces combinatorial issues due to the state space explosion.

In [33], we have presented a new algorithm that exploits the concurrency between transitions of parallel acting components in order to reduce the search space. The algorithm relies on Petri net unfoldings that can be used to compute a compact representation of the dynamics. We have illustrated the applicability of the algorithm with Petri net models of cell signalling and regulation networks, boolean and multi-valued. The proposed approach aims at being complementary to existing methods for deriving the attractors of Boolean models, while being generic since it applies to any safe Petri net.

6.2. Diagnosis

6.2.1. Diagnosability under Weak Fairness

In partially observed Petri nets, diagnosis is the task of detecting whether or not the given sequence of observed labels indicates that some unobservable fault has occurred. Diagnosability is an associated property of the Petri net, stating that in any possible execution an occurrence of a fault can eventually be diagnosed. In [35] we consider diagnosability under the weak fairness (WF) assumption, which intuitively states that no transition from a given set can stay enabled forever; it must eventually either fire or be disabled. Following our previous work [71] on how to perform *weak diagnosis* by exploiting the fact that weak fairness reveals faults in parallel with the current observation, sometimes even before their actual accurrence, we turn to the associated *diagnosability* problem in [35]. First, we show that a previous approach to WF-diagnosability in the literature has a major flaw, and present a corrected notion. Moreover, we present an efficient method for verifying WF-diagnosability based on a reduction to LTL-X model checking. An important advantage of this method is that the LTL-X formula is fixed ? in particular, the WF assumption does not have to be expressed as a part of it (which would make the formula length proportional to the size of the specification), but rather one exploits the ability of existing model checkers to handle weak fairness directly.

6.3. Asynchronous Testing

In the final year of the TECSTES project, we have extended and completed the co-ioco - based conformance and testing theory that we had developed thus far and published in [21], in several directions:

- The testing framework now provides a test generation algorithm [21] for concurrent systems specified with true concurrency models, such as Petri nets or networks of automata. The semantic model of computation of such formalisms are labeled event structures, which allow to represent concurrency explicitly.
- Our test generation algorithm based on Petri net unfolding is able to build a complete test suite w.r.t our co-ioco conformance relation [22]. In addition we propose several coverage criteria that allow to select finite prefixes of an unfolding in order to build manageable test suites.
- We propose an extension of the *ioco* conformance relation, a standard for labeled event structures, named co-ioco, allowing to deal with strong and weak concurrency. We extend the notions of test cases and test execution to labeled event structures, and give a test generation algorithm building a complete test suite for co-ioco. Further, we have introduced and exploited [21] the notions of *strong* and *weak* concurrency: strongly concurrent events must be concurrent in the implementation, while weakly concurrent ones may eventually be ordered, leading to refine *co-ioco* into the *wsc-ioco* relation accounting for weak and strong concurrency.
- The *co-ioco* relation assumes a global control and observation of the system under test, which is not usually realistic in the case of physically distributed systems. Such systems can be partially observed at each of their points of control and observation by the sequences of inputs and outputs exchanged with their environment. Unfortunately, in general, global observation cannot be reconstructed from local ones, so global conformance cannot be decided with local tests. We showed in [39] how appending time stamps to the observable actions of the system under test in order to regain global conformance, via vector clock information, from local testing.
- The MOLE - based testing tool TOURS [42] has been developed with the help of intern Konstantinos Athanasiou, jointly supervised by Hernán Ponce de León and Stefan Schwoon of the MExICo team at LSV), and successful experiments have been conducted with a scalable benchmark example (elevator control). The results show clearly how the true-concurrency approach leads to the test case required being not only smaller individually, but also that *fewer* such test cases are necessary. In addition to the conceptual and analytical enrichment, the results obtained in TECSTES thus also allow to obtain important speedups and reductions in storage space.

Hernán Ponce de León has completed his thesis [40] reporting on the above results, and very successfully defended on Nov. 7, 2014, at ENS Cachan, before the PhD committee consisting of reviewers Rob Hierons and Alex Yakovlev, examiners Thierry Jeron, Remi Morin and Pascal Poizat, and the two supervisors.

6.4. Reachability in MDPs

Markov decision process (MDP) provide the appropriate formalism for the control of fully observable probabilistic systems. There are three kinds of methods for their analysis: linear programming, policy iteration and value iteration. However for large scale systems, only value iteration is still available as it requires less memory than the other methods. For quantitative problems like optimal control for maximizing the discounted reward of an MDP, value iteration is equipped with a stopping criterion that ensures an error bound provided by the user. Value iteration algorithms have also been proposed for the central problem of reachability. However neither stopping criterion nor convergence rate were known for such algorithms. In [37], we have solved these two problems and based on it we have also improved the bound on the number of iterations in order to adapt the value iteration for an exact computation.

6.5. Parameterized Communicating Automata

As a part of our research program on concurrent systems with variable communication topology, we studied system models where the topology is *static* but *unknown*, so that it becomes a parameter of the system. In [28], we introduced parameterized communicating automata (PCAs), where finite-state processes exchange messages via rendez-vous or through bounded FIFO channels. Unlike classical communicating automata, a given PCA can be run on any network topology of bounded degree. We presented various Büchi-Elgot-Trakhtenbrot theorems for PCAs, which roughly read as follows: Let φ be an existential MSO formula and T be any of the following topology classes: pipelines, ranked trees, grids, or rings. There is a PCA that is equivalent to φ on all topologies from T. In the case where each process executes a bounded number of contexts (each context restricting communication in a suitable way), we could show that PCAs are closed under complementation, are expressively equivalent to full MSO logic [29], and have a decidable emptiness problem [31]. The papers [29], [31] are a result of a collaboration with Akshay Kumar (IIT Kanpur) and Jana Schubert (TU Dresden).

6.6. Quantitative behaviours

Several measures have been proposed in literature for quantifying the information leaked by the public outputs of a program with secret inputs. In [32] we studied how to quantify the information leaked by a deterministic or probabilistic program when the measure of information is based on min-entropy or Shannon entropy. A direct computation of these quantities is often infeasible because of the state-explosion problem. In our paper, we model the program as a pushdown system equipped with multi-terminal decision diagrams (ADDs) and propose algorithms to compute said entropies.

The advantage of this approach is that the resulting algorithms can be easily implemented in any BDD-based model-checking tool that checks for reachability in deterministic non-recursive programs by computing program summaries. We demonstrate the validity of our approach by implementing these algorithms in a tool Moped-QLeak.

MIMETIC Project-Team

6. New Results

6.1. Highlights of the Year

6.1.1. Link between performance and risk of injury

Participants: Richard Kulpa [contact], Benoit Bideau, Michaël Ropars.

In our previous biomechanical analysis of the tennis serve, we have demonstrated that the energy flow is a pathomechanical factor, that means that it can increase joint constraints (and thus risk of injury) while not increasing performance. Nevertheless, the definition and evaluation of energy flow is still a complex scientific challenge. We have proposed to compare the energy flow during the serve between injured and non-injured tennis players by investigating the relationships between the quality and magnitude of energy flow, the ball velocity and the peaks of upper limb joint kinetics [11]. The results showed that ball velocity increased and upper limb joint kinetics decreased with the quality of energy flow from the trunk to the 'hand+racket'. Injured players showed a lower quality of energy flow through the upper limb kinetic chain, a lower ball velocity and higher rates of energy absorbed by the shoulder, the elbow than non-injured players. These findings imply that an effective energy flow through the kinetic chain by using a proper serve technique is necessary for reducing overuse joint injury risks.

6.1.2. ACM SIGGRAPH Course on crowd simulation

Participant: Julien Pettré [contact].

Crowds for entertainment or safety applications purposes are most of the time simulated using microscopic algorithms. In contrast with other types of approaches, microscopic simulators are able to generate continuous and smooth trajectories for individual agents. They are based on models of local interactions between agents. The crowd motion result form the combination of all local motion and interactions. The fact that the resulting crowd motion is emergent makes difficult anticipating the simulation results. Many motion and interaction models have been proposed to design a plethora of simulation algorithms: force-based models, rule-based models, coupled or not with flow-based models, etc. Each type of interaction models will actually result into specific crowd motions as well as agents trajectories. Unfortunately, not all have the desired properties: oscillations, jerky trajectories, residual collisions or deadlocks are often observed in simulations. From this point of view, the course [28] presents the many recent progresses in crowd simulation algorithms since the introduction of velocity-based algorithms , as well as the impact on the level of realism and the visual quality of simulated crowd motions. It also presents the impact on various kind of applications.

6.1.3. Immersive basketball playing

Participants: Franck Multon [contact], Alexandra Covaci, Anne-Hélène Olivier.

The paper has received the best paper award of the ACM VRST 2014 Conference in November 2014. This paper addressed the problem of perception of distances in immersive environments when dealing with precision distant tasks, such as basketball free throw. The work has been done in collaboration with University of Brasov in Romania, thanks to the FP7 VISIONAIR infrastructure project. The main results of this work tend to show that third person perspectives enabled subjects to perform the task with similar movements than in real world, compared to first person perspective. Third person perspective consists in placing the camera at a different place from the eye point of view, as in many videogames. On the opposite first person perspective consists in place the camera at the place of the user's eyes in scale 1:1, as if the user was colocalized in the virtual environment. We also demonsrated an adaptation to the task in immersive environments, which is a key information for future development of training methods using VR. We have been invited to submit an extended version of the paper to the IEEE Computer and Graphics journal for 2015.

BEST PAPERS AWARDS :

[21] **ACM Symposium on Virtual Reality Software and Technology VRST**. A. COVACI, A.-H. OLIVIER, F. MULTON.

6.2. VR and Sports

Participants: Richard Kulpa [contact], Benoit Bideau, Franck Multon.

Previous works in MimeTIC have shown the advantage of using VR to design and carry-out experiments on perception-action coupling in sports, especially for duels between two opponents. However the impact of using various technical solutions to carry-out this type of experiment in sports is not clear. Indeed immersion is performed by using interfaces to capture the motion/intention of the user and to deliver various multi-sensory feedbacks. These interfaces may affect the perception-action loop so that results obtained in VR cannot be systematically transferred to real practice.

Most of the applications in VR provide the user with visual feedbacks in which the avatar of the user can be more or less simplified (sometimes limited to a hand or the tools he his carrying). In first person view in caves the user generally does not need accurate avatars as he can perceive his real body but some authors have shown that the perception of distances is generally modified. Some authors have also demonstrated that first person view was less efficient that third person view with avatars when performing accurate tasks such as reaching objects in constrained environments. We proposed an experiment to evaluate which type of feedback was the most appropriate one for complex precision tasks, such as basketball free-throw. In basketball free-throw the user has to throw a ball into a small basket placed at over 4.5m far from him. Thus perception of distance is actually a key point in such a task. Beginners and experts carried-out a first experiment in real in order to measure their motion and performance in real situation. Then beginners were asked to perform free throws with a real ball in hands, but in three conditions in a Cave (Immersia): 1) first person view (see Figure 4), 2) third-person view with the visual feedback of the ball's position, and 3) third-person view the virtual ball and additional rings modeling the perfect trajectory for the ball to get in the basket. Results show that significant difference exist in ball speed between first person view condition compared to real condition whereas no difference exist in third-person view conditions. If we focus on successful throws only, ball speed in the last condition 3) was very similar to real condition whereas all the other VR conditions (1) and 2)) lead to significant differences compared to real situation. In all VR conditions the height of ball release was significantly higher in VR compared to real situation. These results show that VR conditions lead to adaptations in the way people perform such a precision task, especially for ball speed and height of ball release. However this difference is significantly higher with first person view and tends to zero in condition 3). Future works will tend to evaluate new conditions with avatars and complementary points of view (such as lateral and frontal views together as suggested by some authors). It will also be important to more clearly understand the problem of perception of distances in such an environment. This work has been performed in cooperation with University of Brassov in Romania [21]. This paper has received the best paper award of the ACM VRST 2014 Conference.

Another key feedback is the external forces associated with the task. In most sports applications such forces are strongly linked to performance. However delivering these forces in virtual environments is still a challenge as it required haptic devices that could affect the way the users perform the task (with a different grip compared to real situation and limitations in dynamic response of the device). Pseudohaptic has been introduced in the early 2000. It consists in using visual feedbacks to make people perceive the forces linked to a task. However this approach has not been tested for whole-body interaction. In collaboration with Hybrid team in Inria Rennes, we studied how the visual animation of a self-avatar could be artificially modified in real-time in order to generate different haptic perceptions. In our experimental setup participants could watch their self-avatar in a virtual environment in mirror mode. They could map their gestures on the self-animated avatar in real-time using a Kinect. The experimental task consisted in a weight lifting with virtual dumbbells that participants could manipulate by means of a tangible stick. We introduce three kinds of modification of the visual animation of the self-avatar: 1) an amplification (or reduction) of the user motion (change in C/D ratio), 2) a change in the dynamic profile of the motion (temporal animation), or 3) a change in the posture of the avatar (angle of inclination). An example is depicted in Figure 5 . Thus, to simulate the lifting of a "heavy" dumbbell, the

Figure 4. First-person view condition in the basket free-throw.

avatar animation was distorted in real-time using: an amplification of the user motion, a slower dynamics, and a larger angle of inclination of the avatar. We evaluated the potential of each technique using an ordering task with four different virtual weights. Our results show that the ordering task could be well achieved with every technique. The C/D ratio-based technique was found the most efficient. But participants globally appreciated all the different visual effects, and best results could be observed in the combination configuration [9]. Our results pave the way to the exploitation of such novel techniques in various VR applications such as for sport training, exercise games, or industrial training scenarios in single or collaborative mode.

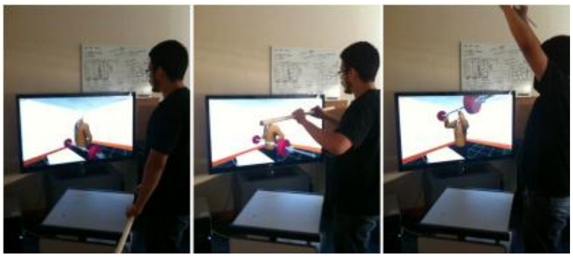


Figure 5. Weight discrimination task: the animation of the avatar showed a lifting effort according to the weight of the virtual dumbbell and the user has to rank the feebacks from the lightest to the heaviest mass.

6.3. Motion Sensing and analysis

Participant: Franck Multon [contact].

Sensing human activity is a very active field of research, with a wide range of applications ranging from entertainment and serious games to personal ambient living assistance, including rehabilitation. MimeTIC aims at proposing original methods to process raw motion capture data in order to compute relevant information according to the application.

In rehabilitation, we have collaborated with University of Montreal, Saint-Justine Hospital which main activity is rehabilitation of children with pathologies of the pyramidal control system. In this domain, defining metrics and relevant measurement to diagnose pathologies and to monitor patients during treatment is a key point. In gait, most of the previous works focus on gait spatio-temporal parameters (such as step length, frequency, stride duration, global speed...) which could be measured with two main families of systems: 1) one-point measurement with a force plate, one accelerometer or dedicated devices (such as a Gait Ride), or 2) multi-point measurement systems with motion sensors or markers placed over the patient's skin. The former provides the clinician with compact but incomplete knowledge whereas the latter provides him with numerous data which are sometimes difficult to analyze and to get (specific technical skills are required). The first step to any type of analysis is to detect the main gait events, such as foot strikes and toe offs. In treadmill walking, widely used in rehabilitation as it enables the clinician to analyze numerous gait cycles in a limited place with a controlled speed, automatically detecting such gait events requires complex devices with specific technical skills (such as calibration and post-processing with motion capture systems) [1].

Recent papers have demonstrated that low-cost and easy-to-use depth cameras (such as a Kinect from Microsoft) look promising for serious applications requiring motion capture. However there exist some confusion between the feet and the ground at foot strike and foot off leading to bad estimation of the gait cycle events. We have proposed an alternative approach that consists in using the strong correlation between knee and foot trajectories to deduce foot strikes thanks to knee movements. Indeed, the maximal distance between knees along the longitudinal axis provides us with very accurate gait events detection compared to previous works. We have validated this detection events on walking patterns that were also altered by placing a 5cm-sole below the left (resp. right) foot of the subject to create asymetry. The results show that this gait cycle event detection based on depth images is as accurate as using reference methods based on accurate motion capture systems.

6.4. VR and Ergonomics

Participants: Charles Pontonnier [contact], Georges Dumont, Pierre Plantard, Franck Multon.

The use of virtual reality tools for ergonomics applications is a very important challenge in order to genezalize the use of such devices for the design of workstations.

We proposed in collaboration with Thierry Duval (Lab-Stic, Telecom Bretagne, Brest) a new architecture for information sharing and bridging in collaborative virtual environements in application to ergonomics studies. We particularly presented this year how we implemented the design engineer role in the collaborative environment [30], [29]. We are currently evaluating the complete framework for collaborative ergonomics by defining use-cases and trying to find the best design mode to efficiently solve this problem. Moreover, we have developed and evaluated some manipulation techniques, such as the 7-handle technique which is particularly efficient to manipulate large objects in an immersive environments [6], [27]. A demonstration of this technique has been presented during the ICAT-EGVE conference [37].

We also contributed in the on-site motion analysis field. Microsoft Kinect is a promising tool to evaluate human poses without markers, calibration and manual post-processing. It has been applied to a wide set of applications, such as entertainment, rehabilitation, sports analysis and more recently in ergonomics. In MimeTIC we wish to develop innovative approaches based on a Kinect in order to assess the potential risks of musculoskeletal disorders. However analyzing humans in work places is challenging because of many potential occlusions and displacements of the user. Hence it is a key point to evaluate to which extent this method could be applied to real work places, in real work condition. Most of previous works aiming at evaluating the Kinect sensor generally focus on simple 2D poses. In this work we proposed to evaluate the reliability of Kinect measurements for assessing the movement of operators in ergonomic studies with complex 3D upper-limb poses [38]. To this end we asked subjects to perform complex 3D arm motions concurrently measured with a Kinect and a Vicon motion capture system. The results demonstrated that most of the poses were correctly estimated with the Kinect but specific poses are badly reconstructed, leaded to errors going up 30°. Hence, experimenter should take this information into account when using a Kinect on a work place in order to avoid experimenting these bad results.

At last, we proposed a new approach for the use of virtual reality with haptics in the Product Development Process loop for testing deformable parts by introducing the user in the loop and proposing a two-stage deformation simulation method for real time haptic interaction. Such an approach is of importance to let the designer be able to handle and validate the design of a product or a workstation respecting multiple constraints, e.g. ergonomics, bulk or productivity. This approach has been fully detailed in a book chapter published this year [34]

6.5. Virtual Human Animation

Participants: Franck Multon [contact], Julien Pettré, Steve Tonneau.

A common issue in three-dimensional animation is the creation of contacts between a virtual creature and the environment. Contacts allow force exertion, which produces motion. This paper addresses the problem of computing contact configurations allowing to perform motion tasks such as getting up from a sofa, pushing an object or climbing. We propose a two-step method to generate contact configurations suitable for such tasks. The first step is an offline sampling of the range of motion (ROM) of a virtual creature. The ROM of the human arms and legs is precisely determined experimentally. The second step is a run time request confronting the samples with the current environment. The best contact configurations are then selected according to a heuristic for task efficiency. The heuristic is inspired by the force transmission ratio. Given a contact configuration, it measures the potential force that can be exerted in a given direction. The contact configurations are then used as inputs for an inverse kinematics solver that will compute the final animation. Our method is automatic and does not require examples or motion capture data. It is suitable for real time applications and applies to arbitrary creatures in arbitrary environments. Various scenarios (such as climbing, crawling, getting up, pushing or pulling objects) are used to demonstrate that our method enhances motion autonomy and interactivity in constrained environments [15], [32]. In Figure 6 , a character is able to select the most appropriate constraints to pull a heavy cupboard by putting a foot on an obstacle to maximize the force ratio.

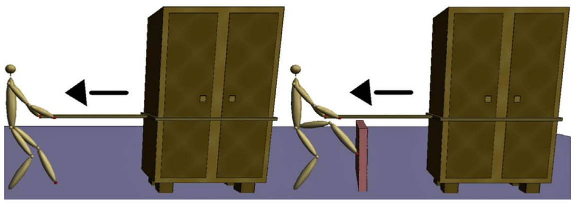


Figure 6. First-person view condition in the basket free-throw.

6.6. Biomechanics for avatar animation

Participants: Charles Pontonnier, Georges Dumont, Steve Tonneau, Franck Multon, Julien Pettré [contact], Ana Lucia Cruz Ruiz, Antoine Muller.

Bio-inpired controllers and planners are compelling for avatar animation. We are currently engaging several works on the subject within the frame of the ENTRACTE project.

Ana-Lucia Cruz-Ruiz has been recruited as a PhD student since november 2013. The goal of this thesis is to define and evaluate muscle-based controllers for avatar animation.A first result has been obtained in defining and validating a bio-inspired limb controller based on a linearizing loop of a neuromuscular complex. Application on a one-dof limb has been validated by comparing the muscle activation shapes obtained in simulation with standard records of biceps and triceps activation [3].

6.7. Semantically consistent herarchical decomposition of virtual urban environments

Participants: Carl-Johan Jorgensen, Fabrice Lamarche [contact].

When planning a path in their environment, pedestrians do not consider every detail at once. Instead, people first plan a coarse path, choosing streets to travel to reach their goal. Local decisions such as where to cross a street or on which side to pass by a pole are taken during navigation. In computer science, hierarchical representations of an environment are often used to reduce the computation cost of the planning algorithm. Such a representation also enables smarter navigation behaviours. Indeed, it others the opportunity to delay the local planning until relevant information is available. It also enables a quick recovery from unexpected events, as the high-level path might stay valid even if unexpected events alter the lower-level path.

We proposed a method that automatically generates a three-level hierarchical representation of an informed urban environment. In this hierarchy, each level is a semantically coherent partition of the navigation areas and can be used to plan paths at different levels of abstraction. This representation is used in a path planning process that delays some decisions until relevant information is perceived. This algorithm uses path options to smartly adapt the path when unexpected events occur.

6.8. Shoulder biomechanics

Participants: Armel Crétual [contact], Michaël Ropars.

At first sight, in clinical practice, shoulder mobility is frequently evaluated through mono-axial amplitude. Interestingly, for diagnosing shoulder hyperlaxity or frozen shoulder, external rotation of the arm whilst at the side (ER1) is commonly used. We first gave a definition of hyperlaxity, as described actually in the literature, and its link with shoulder instability and treatment. After looking for an optimized way to examine external rotation of the shoulder, we proposed the definition of a novel index to quantify global shoulder mobility, the Shoulder Configuration Space Volume (SCSV) corresponding to the reachable volume in the configuration space of the shoulder joint [4]. Then, this new index was examined through correlation to shoulder signs of hyperlaxity [14].

MIMOVE Team

6. New Results

6.1. Introduction

MiMove's research activities in 2014 have focused on a set of areas directly related to the team's research topics. Hence, we have worked on Emergent Middleware (§ 6.3) and Service-oriented Computing in the Future Internet (§ 6.4), in relation to our research topic regarding Emergent Mobile Distributed Systems (§ 3.2). With respect to Large-scale Mobile Sensing & Actuation (§ 3.3), we have developed activities on Service-oriented Middleware for the Mobile Internet of Things (IoT) (§ 6.5), Composing Applications in the IoT (§ 6.6), and Lightweight Streaming Middleware for the IoT (§ 6.7). Last, our effort on Middleware for Mobile Social Networks (§ 6.8) is linked to our research on Mobile Social Crowd-sensing (§ 3.4).

Before presenting our new results in the areas mentioned above, we briefly discuss next the highlights of the year.

6.2. Highlights of the Year

This year has seen the following acknowledgments of the team's contributions:

- Valérie Issarny was distinguished as Chevalier de la Legion d'Honneur for her contributions to science and European scientific cooperation in research and education.
- One of the team's major publication by S. Ben Mokhtar, D. Preuveneers, N. Georgantas, V. Issarny, and Y. Berbers, titled "EASY: Efficient semAntic Service discoverY in pervasive computing environments with QoS and context support" [1], published in the Journal of Systems and Software (Volume 81, Issue 5), is one of the top ten (10) most cited papers among all the papers published by JSS in 2008.

6.3. Emergent Middleware

Participants: Emil Andriescu, Valérie Issarny, Thierry Martinez.

Our previous work on emergent middleware has focused on interconnecting functionally-compatible components, i.e., components that at some high level of abstraction require and provide compatible functionalities, but are unable to interact successfully due to mismatching interfaces and behaviors. To address these differences without changing the components, mediators that systematically enforce interoperability between functionally-compatible components by mapping their interfaces and coordinating their behaviors are required [18]. Our approach for the automated synthesis of mediators is performed through *interface matching*, which identifies the semantic correspondence between the actions required by one component and those provided by the other, followed by the *synthesis of correct-by-construction mediators*. To do so, we analyze the behaviors of components so as to generate the mediator that coordinates the matched actions in a way that guarantees that the two components progress and reach their final states without errors [2]. Our contribution primarily lies in handling interoperability from the application to the middleware layer in an integrated way. The mediators we synthesize act as: (i) translators by ensuring the meaningful exchange of information between components, (ii) controllers by coordinating the behaviors of the components to ensure the absence of errors in their interaction, and (iii) middleware by enabling the interaction of components across the network so that each component receives the data it expects at the right moment and in the right format.

In our latest work, we have particularly focused on item (iii) above. We recognize that modern distributed systems and Systems of Systems (SoS) are built as a composition of existing components and services. As a result, systems communicate (either internally, locally or over networks) using protocol stacks of ever-increasing complexity whose messages need to be translated (i.e., interpreted, generated, analyzed and transformed) by third-party systems. We are particularly interested in the application of message translation to achieve protocol interoperability via protocol mediators. We observe that current approaches are unable to provide an efficient solution towards reusing message translators associated with the message formats composed in protocol stacks. Instead, developers must write ad hoc "glue-code" whenever composing two or more message translators.

Ideally, message translators may be developed by separate parties, using various technologies, while developers should be able to compose them using an easy to use mechanism. However, parsers are monolithic and tightly constructed, which often makes it impossible to combine them, knowing that combining two unambiguous grammars (corresponding to two arbitrary parsers) may result in an ambiguous grammar, and that the ambiguity detection problem for context-free grammars is undecidable in the general case.

In addition to parser composition, the data structures of the parsing output must be manually defined, integrated and harmonized with the target systems (i.e., in this case, the Mediation Engine). As far as we know, the problem of inferring the output schema (or the data type) of an arbitrary tree transformation has not yet been solved, while it is known that, in general, a transformation might not be recognizable by a schema.

Following the challenges above, in [17], we make two major contributions to the issue of systematic message translation for modern distributed systems:

1. Starting from the premise that "off-the-shelf" message translators for individual protocols are readily available in at least an executable form, we propose a solution for the automated composition of message translators. The solution simply requires the specification of a composition rule that is expressed using a subset of the navigational core of the W3C XML query language XPath.
2. We provide a formal mechanism, using tree automata, which based on the aforementioned composition rule, generates an associated AST *data-schema* for the translator composition. This contribution enables the inference of correct data-schemas, relieving developers from the time-consuming task of defining them. On a more general note, the provided method solves the type inference problem for the *substitution* class of tree compositions in linear time on the size of the output. The provided inference algorithm can thus be adapted to a number of applications beyond the scope of this work, such as XML Schema inference for XSLT transformations.

The composition approach that we introduced functions as a purely "black-box" mechanism, thus allowing the use of third-party parsers and message serializers independently of the parsing algorithm they use internally, or the method by which they were implemented/generated. Our solution goes beyond the problem of translator composition by inferring AST data-schemas relative to translator compositions. This feature allows newly generated translators to be seamlessly (or even automatically) integrated with existing systems, and most notably our protocol mediation engine [2].

6.4. Service-oriented Computing in the Future Internet

Participants: Georgios Bouloukakis, Nikolaos Georgantas, Valérie Issarny, Ajay Kattepur, Raphael de Aquino Gomes, Rachit Agarwal.

With an increasing number of services and devices interacting in a decentralized manner, *choreographies* represent a scalable framework for the Future Internet. The service oriented architecture inherent to choreographies allows abstracting diverse systems as application components that interact via standard middleware protocols. However, the heterogeneous nature of such systems leads to choreographies that do not only include conventional services, but also sensor-actuator networks, databases and service feeds. We reason about the behavior of such systems by introducing abstract middleware connectors that follow base interaction paradigms, such as client-service (CS), publish-subscribe (PS) and tuple space (TS). These heterogeneous connectors are made interoperable through a service bus connector, the *eXtensible Service Bus* (XSB) [11].

In previous work, we identified and verified the behavioral semantics of the XSB connector derived from the interconnection of base connectors, and introduced a method for constructing protocol converters enabling this interconnection. We implemented our XSB solution into an extensible development and execution platform for application and middleware designers. We also provided a lightweight implementation of the XSB, the *Light Service Bus* (LSB), appropriate for resource-constrained environments and systems. Next, leveraging on the functional interoperability across interaction paradigms offered by the XSB, we initiated our study of end-to-end Quality of Service (QoS) properties of choreographies, where in particular we focus on the effect of middleware interactions on QoS.

Building on the above results, we refine our analysis of QoS on top of the identified interaction paradigms. We have introduced a motivating application scenario inspired from the *2014 D4D Challenge*[0]. More specifically, *Data for Development Senegal* is an innovation challenge on ICT Big Data for the purposes of societal development. Mobile network provider Sonatel (part of the Orange Group) has made anonymous data extracted from the mobile network in Senegal available to international research laboratories, encouraging research related to the development and welfare of the local population.

Our scenario targets the development of an application platform for citywide and countrywide transport information management relying on mobile social crowd-sensing. This takes into account the particular context and constraints in Senegal. More specifically, the local transportation system, although developing, still consists of many unplanned and informal settlements with unreliable services and infrastructure. Additionally, despite wide use of mobile phones in the country, mobile Internet access remains limited, making SMS the only alternative for data access for a large part of the population. Our proposition aims to complement the scarce authoritative transport information coming from structured information sources and compensate for the lack of such information. In particular, in our approach we intend to study and experiment with appropriate interaction paradigms (CS, PS, TS) on top of 3G/2G/SMS data connections, further depending on the specific application and data. We are especially interested in interaction adaptation depending on the network conditions (e.g., switching to SMS-based protocol when the 3G/2G network is unavailable).

We have taken a first step towards enabling such an application platform. This consists in evaluating the publish/subscribe interaction style in a large-scale setting where resources of mobile users are limited, which translates into limited and intermittent connectivity in the system. Additionally, such an application platform must guarantee that the sensing data is processed and delivered to the corresponding mobile users *on-time*, despite the intermittent connectivity of the latter. We have opted for the publish/subscribe paradigm, as it is deemed appropriate for loose spatio-temporal interaction between mobile entities.

In particular, we introduce a queueing network model for the end-to-end interaction within a large-scale mobile publish/subscribe system. We leverage the *D4D dataset* provided by Orange Labs to parametrize this model. We then develop a simulator named *MobileJINQS*[0] that implements our model and uses the dataset traces as realistic input load to the system model over the time span of a whole year. Prior to this, we extensively analyze the D4D dataset in order to identify the data that we are interested in and infer primary results [0]. Based on the results of our simulation-based experiments, we thoroughly evaluate the behavior of the publish/subscribe system and identify ways of tuning the system parameters in order to satisfy certain design requirements. More precisely, we provide results of simulations of our publish/subscribe system with varied incoming loads, service delays and event lifetime periods. We use connection data of various pairs of mobile network antennas to derive realistic traces for both incoming loads and service delays. System or application designers are able to tune the system by selecting appropriate lifetime periods. We demonstrate that varying incoming loads and service delays have a significant effect on response time. By properly setting event lifetime spans, designers can best deal with the tradeoff between freshness of information and information delivery success rates. Still, both of these properties are highly dependent on the dynamic correlation of the event input flow and delivery flow processes, which are intrinsically decoupled.

[0] http://www.d4d.orange.com/en/home
[0] http://xsb.inria.fr/d4d#mobilejinqs
[0] http://xsb.inria.fr/d4d

Our future work includes comparison of the publish/subscribe interaction paradigm with other interaction paradigms (client-server, tuple space), in relation with the network access capacity and the application requirements. Also, we intend to study the response time and success rate for the various combinations of antennas in more fine-grained scales (e.g., check what their evolution is over one day).

6.5. Service-oriented Middleware for the Mobile Internet of Things

Participants: Sara Hachem, Valérie Issarny, Georgios Mathioudakis, Animesh Pathak, Fadwa Rebhi.

The Internet of Things (IoT) is characterized by a wide penetration in the regular user's life through an increasing number of Things embedding sensing, actuating, processing, and communication capacities. A considerable portion of those Things will be mobile Things, which come with several advantages yet lead to unprecedented challenges. The most critical challenges, that are directly inherited from, yet amplify, today's Internet issues, lie in handling i) the large scale of users and mobile Things which lead to high communication and computation costs especially with the anticipated large volumes of data to exchange, ii) providing interoperability across the heterogeneous Things which host sensors and actuators providing services and producing data that follow different format/schema specifications, and iii) overcoming the unknown dynamic nature of the environment, due to the mobility of an ultra-large number of Things.

Service-Oriented Architecture (SOA) provides solid basis to address the above challenges as it allows the functionalities of sensors/actuators embedded in Things to be provided as services, while ensuring loose-coupling between those services and their hosts, thus abstracting their heterogeneous nature. In spite of its benefits, SOA has not been designed to address the ultra-large scale of the mobile IoT. Consequently, an alternative is provided within a novel Thing-based Service-Oriented Architecture, that revisits SOA interactions and functionalities, service discovery and composition in particular. Our work on the revisited Thing-based SOA is detailed in [9], [23], [15]. The novel architecture is concretized within MobIoT, a middleware solution that is specifically designed to manage and control the ultra-large number of mobile Things in partaking in IoT-related tasks.

In accordance with SOA, MobIoT comprises Discovery, Composition & Estimation, and Access components, yet modifies their internal functionalities. In more detail, the Discovery component enables Thing-based service registration (for Things to advertise hosted services) and look-up (for Things to retrieve remote services of interest). In order to handle the ultra large number of mobile Things and their services in the IoT, the component revisits the Service-Oriented discovery and introduces probabilistic protocols to provide, not all, but only a sufficient subset of services that can best approximate the result that is being sought after [23], [15] based on a predefined set of requirements such us sensing coverage of the area of interest and the location of the Things. By limiting the participation of Things, the communication costs and volumes of data to process are decreased without jeopardising the quality of the outcome.

Furthermore, the Composition & Estimation component (C&E) provides automatic composition of Thing-based services. This capacity is of interest in the case where no service can perform a required measurement/action task directly (based on its atomic functionalities). To that end, we model our composition specification as mathematical formulas defined semantically within a dedicated ontology. Thing-based service composition executes in three phases: i) expansion, where composition specifications are automatically identified; ii) mapping, where actual service instances (running services) are selected based on their functionalities and the physical attributes of their hosts; and iii) execution, where the services are accessed and the composition specifications are executed. Thing-based service composition revisits Service-Oriented composition by executing seamlessly with no involvement from developers or end users and relying on semantic technologies to identify the most appropriate services to compose.

Last but not least, the Access component provides an easy to use interface for developers to sample sensors/actuators while abstracting sensor/actuator hardware specifications. It revisits Service-Oriented access and leverage semantic technologies by executing access to services transparently and wrapping access functionalities internally. Thus, it alleviates that burden from users, initially in charge of this task. The Access component supports real-time query-based access to remote services and to locally hosted services.

To assess the validity of our proposed architecture, we provide a prototype implementation of MobIoT(§ 5.4) along with a set of extensive evaluations that demonstrate, not only the feasibility of our approach, but also the resulting quality of the discovery approach, along with its scalability, as compared to a regular SOA-based approach.

6.6. Composing Applications in the Internet of Things

Participants: Aness Bajia, Animesh Pathak, Françoise Sailhan.

Resilient computing is defined as the ability of a system to remain dependable when facing changes. To mitigate faults at runtime, dependable systems embark fault tolerance mechanisms such as replication techniques. These mechanisms have to be systematically and rigorously applied in order to guarantee the conformance between the application runtime behavior and its dependability requirements.

Given that devices and networks constituting the IoT are prone to failure and consequent loss of performance, it is natural that IoT applications are expected to encounter and tolerate several classes of faults - something that still largely remains within the purview of low-level-protocol designers. As part of our work on the MURPHY project (§ 7.1.1.1), we are addressing this issue by proposing: i) a set of abstractions that can be used during macroprogramming to express application-level fault tolerance requirements, as well as by developers of fault tolerance protocols to identify the abilities and requirements of their techniques; ii) a runtime system that employs adaptive fault tolerance (AFT) to provide fault tolerance to the networking sensing application; and iii) compilation techniques to instantiate and map tasks as needed to satisfy the requirements of the application for a given deployment. Through our work [26], we demonstrate that our approach provides this much-needed feature to networked sensing applications with negligible development- and minimal performance- overhead.

Complementary to the above, we have proposed task mapping algorithms to satisfy those requirements through a constraint programming approach [24]. Through evaluations on realistic application task graphs, we show that our constraint programming model can effectively capture the end-to-end requirements and efficiently solves the combinatorial problem introduced.

We have been continually incorporating our research results in the above areas into *Srijan* (§ 5.5), which provides an easy-to-use graphical front-end to the various steps involved in developing an application using the ATaG macroprogramming framework.

6.7. Lightweight Streaming Middleware for the Internet of Things

Participants: Benjamin Billet, Valérie Issarny.

The IoT raises many challenges related to its very large scale and high dynamicity, as well as the great heterogeneity of the data and systems involved (e.g., powerful versus resource-constrained devices, mobile versus fixed devices, continuously-powered versus battery-powered devices, etc.). These challenges require new systems and techniques for developing applications that are able to: (i) collect data from the numerous data sources of the IoT, and (ii) interact both with the environment using the actuators and with the users using dedicated GUIs. Given the huge volume of data continuously being produced by sensors (measurements and events), we must consider: (i) data streams as the reference data model for the IoT and, (ii) continuous processing as the reference computation model for processing these data streams. Moreover, knowing that privacy preservation and energy consumption are increasingly critical concerns, we claim that all the Things should be autonomous and work together in restricted areas as close as possible to the users rather than systematically shifting the computation logic into powerful servers or into the cloud.

Toward that goal, we have been developing Dioptase [3], a service-oriented middleware for the IoT, which aims to integrate the Things and their streams into today's Web by presenting sensors and actuators as Web services. The research work around the Dioptase middleware consists in designing new service-oriented architectures where services continuously process data streams instead of finite datasets. In this context, new composition mechanisms are investigated in order to provide a way to describe complex fully-distributed stream-based tasks and to deploy them dynamically, at any time, as task graphs, over available Things of the network, including resource-constrained ones. To this end, Dioptase enables task graphs to be composed of Thing-specific tasks (directly implemented on the Thing) and dynamic tasks that communicate using data streams. Dynamic tasks are then described in a lightweight DSL, called *DiSPL*, which is directly interpreted by the middleware and provides specific primitives to manipulate data streams.

As part of the design of such composition mechanisms, we have been investigating the problem of task mapping and automated deployment, which basically consists of mapping a set of tasks onto a set of nodes. Given the specific challenges introduced by the IoT, we worked on a new formalization of the task mapping problem that captures the varying consumption of resources and various constraints (location, capabilities, QoS) in order to compute a mapping that guarantees the lifetime of the concurrent tasks inside the network and the fair allocation of tasks among the nodes (load balancing). This formalization, called *Task Graph to Concrete Actions (TGCA)* [19], results in a binary programming problem for which we provide an efficient heuristic that allows its resolution in polynomial time. Our experiments show that our heuristic: (i) gives solutions that are close to optimal, and (ii) can be implemented on reasonably powerful Things and performed directly within the network without requiring any centralized infrastructure.

6.8. Middleware for Mobile Social Networks

Participants: Animesh Pathak, George Rosca.

As recent trends show, online social networks (OSNs) are increasingly turning mobile and further calling for decentralized social data management. This trend is only going to increase in the near future, based on the increased activity, both by established players like Facebook and new players in the domain such as Google, Instagram, and Pinterest. Modern smart phones can thus be regarded as *social sensors*, collecting data not only passively using, e.g., Bluetooth neighborhoods, but actively in the form of, e.g., "check-in"s by users to locations. The resulting (mobile) social ecosystems are thus an emergent area of interest.

The recent years have seen three major trends in the world of online social networks: *i)* users have begun to care more about the privacy of their data stored by large OSNs such as Facebook, and have won the right (at least in the EU) to remove it completely from the OSN if they want to; *ii)* OSNs are making their presence felt beyond casual, personal interactions to corporate, professional ones as well, starting with LinkedIn, and most recently with the purchase by Microsoft of Yammer, the enterprise social networking startup, and the launch of Google Plus for enterprise customers; and *iii)* users are increasingly using the capabilities of their (multiple) mobile devices to enrich their social interactions, ranging from posting cellphone-camera photos on Instagram to "checking-in" to a GPS location using Foursquare.

In view of the above, we envision that in the near future, the use of ICT to enrich our social interactions will grow (including both personal and professional interactions), both in terms of size and complexity. However current OSNs act mostly like data silos, storing and analyzing their users' data, while locking in these users to their servers, with non-existent support for federation; this is reminiscent of the early days of email, where one could only email those who had accounts on the same Unix machine. The knee-jerk reaction to this has been to explore completely decentralized social networks, which give the user complete control over and responsibility of their social data, while resorting to peer-to-peer communication protocols to navigate their social networks. Unfortunately, there are few techniques available to reconcile with the fact that the same user might have multiple devices, or that it is extremely resource-consuming to perform complex analysis of social graphs on small mobile devices.

Our view lies somewhere in the middle of the two extremes, taking inspiration from the manner in which users currently use email. While their inboxes contain an immense amount of extremely personal data, most users are happy to entrust it to corporate or personal email providers (or store and manage it individually on their personal email servers) all the while being able to communicate with users on any other email server. The notion of *Federated Social Networks* (FSNs) —already gaining some traction— envisions a similar ecosystem where users are free to choose OSN providers which will provide storage and management of their social information, while allowing customers using different OSN providers to interact socially. Such a federation can be beneficial in three major ways, among others: *i)* it allows users to enjoy properties such as reliability, availability, and computational power of the hosting infrastructure of their choice, while not being locked down in terms of whom they can communicate with; *ii)* much like spam filtering services provided by modern email providers, that are tuned by feedback from their users, FSN users can benefit from the behavior of others sharing the same OSN provider [0]; and *iii)* this fits perfectly with enterprise needs, where ad-hoc teams can be formed across corporate OSN providers of two organizations to work on a joint project.

In [30], we presented a set of requirements, followed by a survey of the state of the art in social networking solutions, with a special focus on their ability to support rich privacy and access control policies in federated settings. Through this extensive analysis we offer a broad vision on existing social networking platforms, protocols involved but also their privacy and access policies. By doing so, we identify the main components of a federated social platform together with presenting the current trends in standards and security paradigms underlying actual open source solutions which offers their implementation, and finally provides recommendations on constructing such systems. Our research is continually being incorporated into the Yarta middleware for mobile social networking(§ 5.7).

[0] This also gives an incentive to commercial OSN providers to provide value-added services.

MINT Project-Team

6. New Results

6.1. Highlights of the Year

- "Adoiraccourcix : sélection de commandes sur écrans tactiles multi-points par identification des doigts" [31] received the *best paper award* from the IHM 2014 conference;
- "L'ordinateur portable comme instrument de musique" [41] received the *best demo award* from the IHM 2014 conference.

6.2. Impact of form factors and input conditions on absolute indirect-touch pointing tasks

Absolute indirect interaction maps the absolute position of a device's end-effector to the absolute position of a remote on screen object. Despite its long-time use with graphics tablets and growing use in research prototypes, little is known on the influence of form factors and input conditions on pointing performance with such a mapping. The input and display can have different sizes and aspect ratios, for example. The on-screen targets can vary in size. Users can look solely at the display or at the input device as well. They can also hold the input device in certain cases, or let it rest on a table. We ran two experiments designed to investigate the influence of all these factors on absolute indirect-touch pointing performance [20], [11].

The first experiment focused on input device size and input conditions and revealed that users get higher performance when they can look at the input surface (even if nothing is displayed on it). In addition we found that the smallest target size users can acquire in motor space is not constant across different input dimensions but degrades as the input size increases. The second experiment focused on scale effects and aspect ratio and revealed users' performance is not affected by scale but that aspect ratio matters: similar input and output aspect ratios lead to better performance. This findings led us to list four main recommendations for the design of touch input surfaces with applications supporting absolute direct interaction.

6.3. Direct and indirect multi-touch interaction on a wall display

Multi-touch wall displays allow to take advantage of co-located interaction (direct interaction) on very large surfaces. However interacting with content beyond arms' reach requires body movements, introducing fatigue and impacting performance. Interacting with distant content using a pointer can alleviate these problems but introduces legibility issues and loses the benefits of multi-touch interaction. We introduced WallPad [30], [11], a widget designed to quickly access remote content on wall displays while addressing legibility issues and supporting direct multi-touch interaction (Figure 1). To support multi-touch on such a wall display, we developed a custom system using front diffuse illumination and 4 cameras. Our system can detect 50+ simultaneous contacts with a precision between 3 and 5 mm.

6.4. Sketching dynamic and interactive illustrations

We collaborated with Autodesk Research in Toronto (as a scientific consultant) on a project whose focus was to design and develop tools that enable artists to bring life to illustrations with subtle, continuous animation effects and infusing interactive behavior to the drawings. We believe designers, artists and creators should be able to communicate with computers the way they think about art and animation. This motivated Autodesk to develop interfaces that facilitate powerful ways of thinking and content creation with freeform sketching and direct manipulation, thus offering an alternative to complex professional animation tools. Our design combines the complementary affordances of humans and computers by utilizing by-example phenomena, thus preserving expressiveness and personal style, yet reducing tedium.

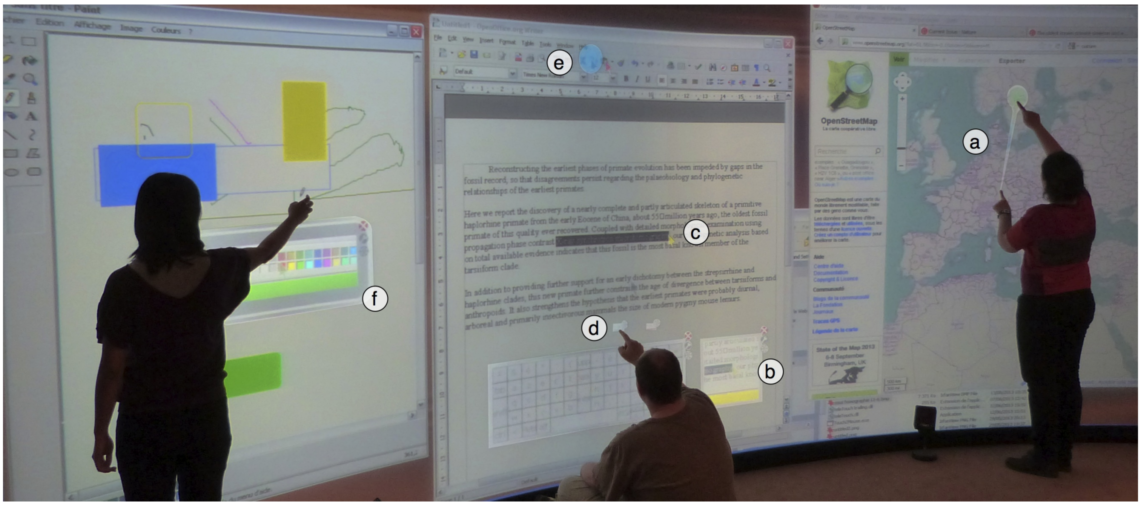


Figure 1. A user creating a WallPad (a). Another using two WallPads to select (b) some text (c) and click (d) on Copy/Paste buttons (e) beyond reach. A third user accessing a color palette located at the bottom of the screen through a WallPad (f).

The outcome of the collaboration is Kitty [23], a sketch-based tool for authoring dynamic and interactive illustrations (Figure 2). Artists can sketch animated drawings and textures to convey the living phenomena, and specify the functional relationship between its entities to characterize the dynamic behavior of systems and environments. An underlying graph model, customizable through sketching, captures the functional relationships between the visual, spatial, temporal or quantitative parameters of its entities. As the viewer interacts with the resulting dynamic interactive illustration, the parameters of the drawing change accordingly, depicting the dynamics and chain of causal effects within a scene. The generality of this framework makes our tool applicable for a variety of purposes, including technical illustrations, scientific explanation, infographics, medical illustrations, children's e-books, cartoon strips and beyond. A user study demonstrates the ease of usage, variety of applications, artistic expressiveness and creative possibilities of our tool.

Kitty is a follow up of a previous project, Draco [50], a prototype sketch-based interface that allows artists and casual users alike to add a rich set of animation effects to their drawing, seemingly bringing illustrations to life such as a school of fish swimming, tree leaves blowing in the wind, or water rippling in a pond. Draco was realized before Fanny Chevalier joined Inria. Kitty is the result of a collaboration between Autodesk Research (inventor) and Inria (scientific consultant). A patent has been filed by Autodesk Research for Kitty, and the company is currently developping a commercial application based on the research protoype.

6.5. The not-so-staggering effect of staggered animations

Interactive visual applications often rely on animation to transition from one display state to another. There are multiple animation techniques to choose from, and it is not always clear which should produce the best visual correspondences between display elements. One major factor is whether the animation relies on staggering—an incremental delay in start times across the moving elements. It has been suggested that staggering may reduce occlusion, while also reducing display complexity and producing less overwhelming animations, though no empirical evidence has demonstrated these advantages. We empirically evaluated the effect of two staggering techniques on tracking tasks, focusing on cases that should most favour staggering [14]. We found that introducing staggering has a negligible, or even negative, impact on multiple object tracking performance. The potential benefits of staggering may be outweighed by strong costs: a loss of common-motion grouping

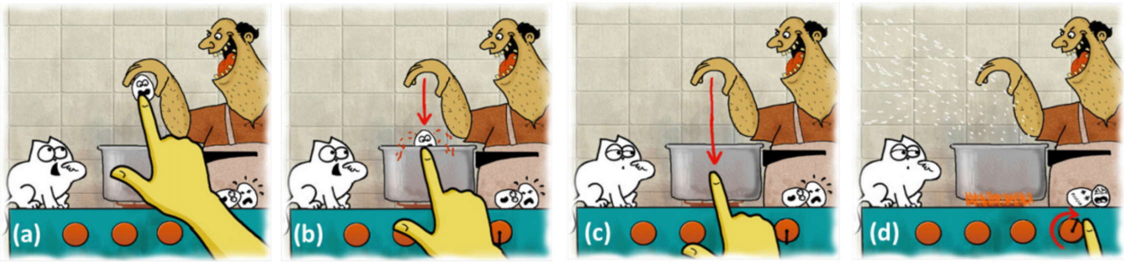


Figure 2. Example of a dynamic interactive illustration authored with Kitty. (a) Objects in the scene are interactive: the egg held by the cook can be dragged down, as if falling into the pot, triggering subsequent animations, such as soup splashes (b) and closing of the cat's eyelids (c). Turning the knob increases the fire and steam (d). The resulting dynamic illustration captures the living nature of the scene, where the gas stove flames burn and steam emits from the pot.

information about which objects travel in similar paths, and less predictability about when any specific object would begin to move.

6.6. Flexible contextual retrieval of chosen documents and windows

Users of Personal Computers interact with a large number of resources to do their work. To handle their different tasks, they need their documents to be readily available, and as the number of activities and documents increase, systems must offer adequate support for quick retrieval of these resources. The Hotkey Palette [29] is a quick retrieval facility that we designed that uses hotkeys and makes them visible and configurable through a quasi-modal always-available on-screen keyboard. This facility contributes to the state of the art in three ways. It extends on-screen keyboard interaction by providing feedback on the state of the linked resources, it provides persistent and integrated access to local windows and files and other online resources, and it provides flexible control over contextualization by leveraging existing resource hierarchies.

6.7. Multi-touch command selection using finger identification

Hotkeys are a critical factor of performance for expert users in WIMP interfaces. Multi-touch interfaces, by contrast, do not provide such efficient command shortcuts. Adoiraccourcix leverages finger identification to introduce quick command invocation integrated with direct manipulation in this context (Figure 3). We illustrated its use in a vectorial drawing application and ran preliminary user studies comparing it to classical user interfaces. Results suggest that once mastered, it provides very powerful means of interaction [31], [44], [43].

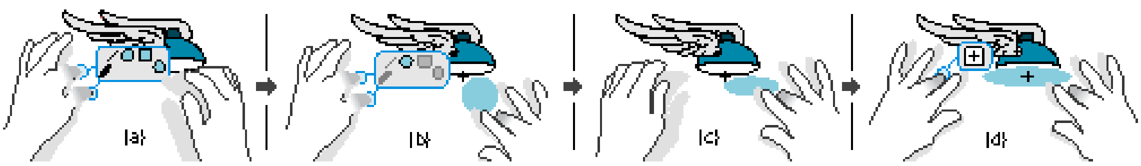


Figure 3. Partial illustration of the Adoiraccourcix' logics.

6.8. Impact of the localization and activation of mode switchers

Input devices have a limited number of buttons and degrees of freedom, but they are used to control many functionnalities. Modes and quasi-modes makes it possible to map several actions to the same input. For example keys of a keyboard either input a letter or trigger a command. Delimiters allow users to switch between the modes. On the keyboard, the default mode is often text entry and pressing the Ctrl key switches the mode to command mode. This choice was made at a time when the mouse was not widespread. In [33] , we explored the possibility to place mode switchers on the mouse and experimented the benefits. We showed that there is a performance benefit if the current tasks are essentially mouse-based. In particular we showed that using mode switchers on the mouse reduces homing the dominant hand between the mouse and the keyboard.

6.9. A serial Architecture for a collaborative robot

The haptic magnifier consists in using a serial architecture, where a motor is inserted between a tool and a user's hand (figure 4). By this way, the tool's speed v_o can be changed relatively to user's speed v_i, by controlling motor's speed. The haptic rendering of a load can then be changed, and fine details can be more easily detected.

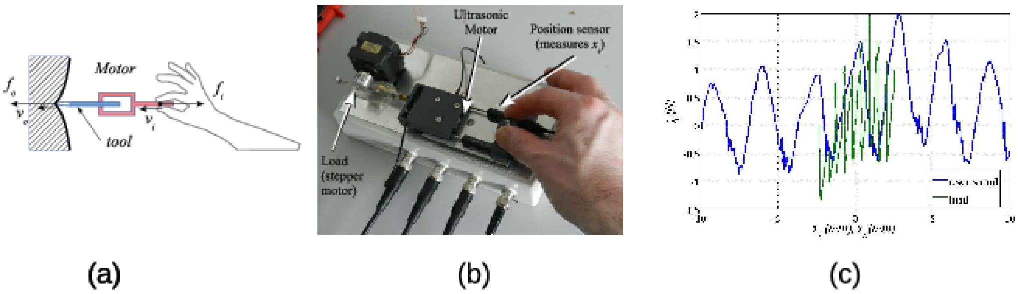


Figure 4. The Haptic Magnifier; (a) the serial architecture with a motor inserted to achieve a haptic magnifier, (b) the implementation with an ultrasonic Motor, and (c) the resulting rendering at load's end and user's end.

The haptic magnifier is built up with an ultrasonic motor, whose characteristic is low speed - high torque. So the tool and the end-effector can be directly connected to the motor, leading to a lightweight architecture. The user's study presented in [21] have shown that the precision in using the tool could be improved during a freehand manipulation.

6.10. Mimetic Interaction Spaces: Controlling Distant Displays in Pervasive Environments

Pervasive computing is a vision that has been an inspiring long-term target for many years now. Interaction techniques that allow one user to efficiently control many screens, or that allow several users to collaborate on one distant screen, are still hot topics, and are often considered as two different questions. Standard approaches require a strong coupling between the physical location of input device, and users. We propose to consider these two questions through the same basic concept, that uncouples physical location and user input, using a mid-air approach. We present the concept of mimetic interaction spaces (MIS), a dynamic user-definition of an imaginary input space thanks to an iconic gesture, that can be used to define mid-air interaction techniques. We describe a participative design user-study, that shows this technique has interesting acceptability and elicit some definition and deletion gestures. We finally describe a design space for MIS-based interaction, and show how such concept may be used for multi-screen control, as well as screen sharing in pervasive environments [26].

6.11. Match-Up & Conquer: A Two-Step Technique for Recognizing Unconstrained Bimanual and Multi-Finger Touch Input

We present a simple, two-step technique for recognizing multi-touch gesture input independently of how users articulate gestures, i.e., using one or two hands, one or multiple fingers, synchronous or asynchronous stroke input. To this end, and for the first time in the gesture literature, we introduce a preprocessing step specifically for multi-touch gestures (Match-Up) that clusters together similar strokes produced by different fingers, before running a gesture recognizer (Conquer). We report gains in recognition accuracy of up to 10% leveraged by our new preprocessing step, which manages to construct a more adequate representation for multi-touch gestures in terms of key strokes. It is our hope that the Match-Up technique will add to the practitioners toolkit of gesture preprocessing techniques, as a first step toward filling todays lack of algorithmic knowledge to process multi-touch input and leading toward the design of more efficient and accurate recognizers for touch surfaces. [27]

6.12. Understanding Users's perceived Difficulty of Multi-Touch Gesture Articulation

We show that users are consistent in their assessments of the articulation difficulty of multi-touch gestures, even under the many degrees of freedom afforded by multi-touch input, such as (1) various number of fingers touching the surface, (2) various number of strokes that structure the gesture shape, and (3) single-handed and bimanual input. To understand more about perceived difficulty, we characterize gesture articulations captured under these conditions with geometric and kinematic descriptors computed on a dataset of 7,200 samples of 30 distinct gesture types collected from 18 participants. We correlate the values of the objective descriptors with users' subjective assessments of articulation difficulty and report path length, production time, and gesture size as the highest cor- relators (max Pearson's r=.95). We also report new findings about multi-touch gesture input, e.g., gestures produced with more fingers are larger in size and take more time to produce than single-touch gestures; bimanual articulations are not only faster than single-handed input, but they are also longer in path length, present more strokes, and result in gesture shapes that are deformed horizontally by 35% in average. We use our findings to outline a number of 14 guidelines to assist multi-touch gesture set design, recognizer development, and inform gesture-to-function mappings through the prism of the user-perceived difficulty of gesture articulation.[28]

6.13. Dynamic Modelling of Electrovibration

Electrostatic attraction may be used to modulate the apparent friction coefficient between a fingertip and a surface to create a tactile stimulator. In this work, we want to propose an accurate modelling of the force generation. For that purpose, a specific experimental test bench has been manufactured, as presented in figure 5 .

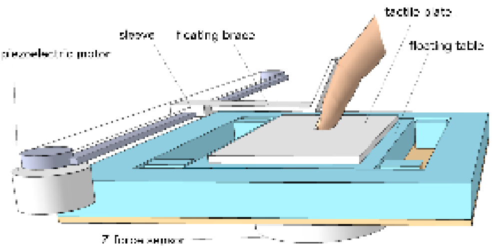


Figure 5. Representation of the measurement system, the finger is moved on the plate by the motor.

Then, an investigation on the current modeling were carried out, with a focus on the temporal evolution and frequency dependence of the stimulus. More particularly, we considered the charge lost through the stratum corneum. Indeed, lost charges is gathered on the surface of the insulator as free surface charge, for this reason it no longer participates to the generation of the force on the finger, and consequently, to the measured force (Fig. 6). This happens because the charges on the surface of the insulator are no longer mechanically bounded to the finger and the insulator sustains the induced electrostatic force.

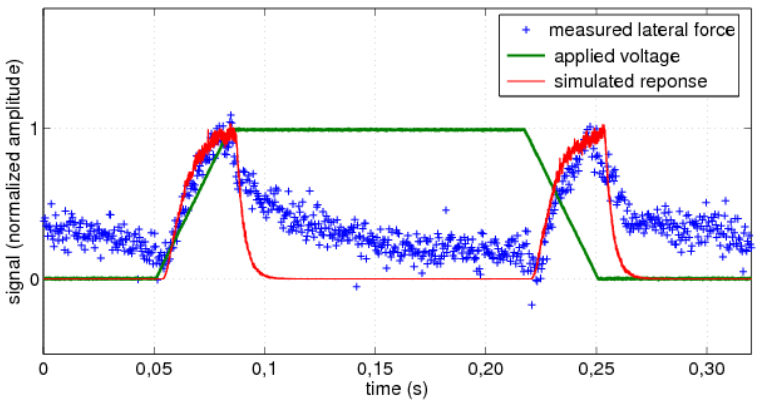


Figure 6. Charge configuration at the border of the stratum corneum (SC) and insulator (I). The conductive part of the system is represented like the electrode of a capacitor. (1) Initial configuration on the charge when the voltage v is applied. (2) Discharge through the stratum corneum with the two equivalent capacitors. (3) Final configuration of the charges after the transient.

The improvement of the modeling is proposed to take into account this major effect, and then, it is checked with an experimental set-up and compared with literature results.

6.14. Coupling between Electrovibration and squeeze filmfor tactile stimulation

Electrovibration and squeeze film effect are two different principles which modify user perception of a surface. The first is generated by a polarization of a finger approaching a high voltage supplied plate, and the latter by an ultrasonic vibrating plate. Their compatibility on the same stimulator has been analysed and their concomitant has been proven as well as the increased range of sensations [34]. A joint model has been proposed to describe the behaviour of the friction when both principles are merged. For the analysis, a specific experimental test bench has been built to measure the forces induced, as shown in figure 7 .

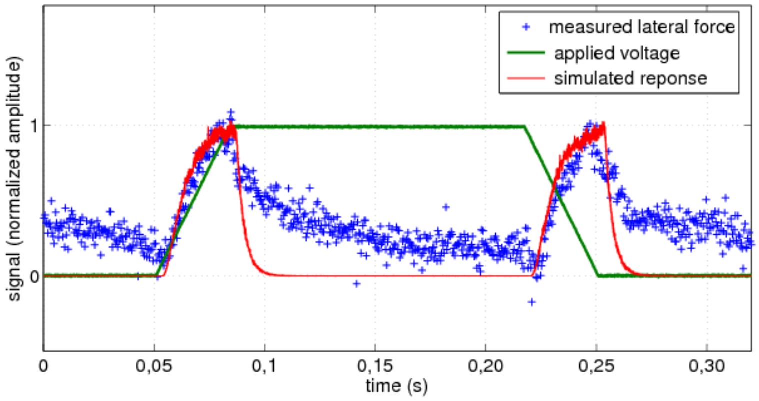


Figure 7. The experimental setup, and the recorded friction modulation.

MISTIS Project-Team

6. New Results

6.1. Highlights of the Year

6.1.1. P-Locus software and Pixyl start-up project

The work on the P-Locus software has been exploited in order to create a start-up in January 2015. The project called Pixyl have been accepted by the GATE1 incubator and has been awarded a BPI emergence prize. It is leaded by Senan Doyle (future CEO). The other co-founders are Michel Dojat (INSERM, GIN), Florence Forbes (Inria, Mistis) and IT-Translation.

6.2. Mixture models

6.2.1. Parameter estimation in the heterogeneity linear mixed model

Participant: Marie-José Martinez.

Joint work with: Emma Holian (National University of Ireland, Galway)

In studies where subjects contribute more than one observation, such as in longitudinal studies, linear mixed models have become one of the most used techniques to take into account the correlation between these observations. By introducing random effects, mixed models allow the within-subject correlation and the variability of the response among the different subjects to be taken into account. However, such models are based on a normality assumption for the random effects and reflect the prior belief of homogeneity among all the subjects. To relax this strong assumption, Verbeke and Lesaffre (1996) proposed the extension of the classical linear mixed model by allowing the random effects to be sampled from a finite mixture of normal distributions with common covariance matrix. This extension naturally arises from the prior belief of the presence of unobserved heterogeneity in the random effects population. The model is therefore called the heterogeneity linear mixed model. Note that this model does not only extend the assumption about the random effects distribution, indeed, each component of the mixture can be considered as a cluster containing a proportion of the total population. Thus, this model is also suitable for classification purposes.

Concerning parameter estimation in the heterogeneity model, the use of the EM-algorithm, which takes into account the incomplete structure of the data, has been considered in the literature. Unfortunately, the M-step in the estimation process is not available in analytic form and a numerical maximisation procedure such as Newton-Raphson is needed. Because deriving such a procedure is a non-trivial task, Komarek et al. (2002) proposed an approximate optimization. But this procedure proved to be very slow and limited to small samples due to requiring manipulation of very large matrices and prohibitive computation.

To overcome this problem, we have proposed in an alternative approach which consists of fitting directly an equivalent mixture of linear mixed models. Contrary to the heterogeneity model, the M-step of the EM-algorithm is tractable analytically in this case. Then, from the obtained parameter estimates, we can easily obtain the parameter estimates in the heterogeneity model.

6.2.2. Taking into account the curse of dimensionality

Participants: Stéphane Girard, Alessandro Chiancone, Seydou-Nourou Sylla.

Joint work with: C. Bouveyron (Univ. Paris 5), M. Fauvel (ENSAT Toulouse) and J. Chanussot (Gipsa-lab and Grenoble-INP)

In the PhD work of Charles Bouveyron (co-advised by Cordelia Schmid from the Inria LEAR team) [67], we propose new Gaussian models of high dimensional data for classification purposes. We assume that the data live in several groups located in subspaces of lower dimensions. Two different strategies arise:

- the introduction in the model of a dimension reduction constraint for each group
- the use of parsimonious models obtained by imposing to different groups to share the same values of some parameters

This modelling yields a new supervised classification method called High Dimensional Discriminant Analysis (HDDA) [4]. Some versions of this method have been tested on the supervised classification of objects in images. This approach has been adapted to the unsupervised classification framework, and the related method is named High Dimensional Data Clustering (HDDC) [3]. Our recent work consists in adding a kernel in the previous methods to deal with nonlinear data classification and heterogeneous data [12]. We also investigate the use of kernels derived from similary measures on binary data. The targeted application is the analysis of verbal autopsy data (PhD thesis of N. Sylla): Indeed, health monitoring and evaluation make more and more use of data on causes of death from verbal autopsies in countries which do not keep records of civil status or with incomplete records. The application of verbal autopsy method allows to discover probable cause of death. Verbal autopsy has become the main source of information on causes of death in these populations.

6.2.3. Location and scale mixtures of Gaussians with flexible tail behaviour: properties, inference and application to multivariate clustering

Participant: Florence Forbes.

Joint work with: Darren Wraith from QUT, Brisbane Australia.

Clustering concerns the assignment of each of N, possibly multidimensional, observations $y_1, ..., y_N$ to one of K groups. A popular way to approach this task is via a parametric finite mixture model. While the vast majority of the work on such mixtures has been based on Gaussian mixture models in many applications the tails of normal distributions are shorter than appropriate or parameter estimations are affected by atypical observations (outliers). The family of location and scale mixtures of Gaussians has the ability to generate a number of flexible distributional forms. It nests as particular cases several important asymmetric distributions like the Generalised Hyperbolic (GH) distribution. The Generalised Hyperbolic distribution in turn nests many other well known distributions such as the Normal Inverse Gaussian (NIG) whose practical relevance has been widely documented in the literature. In a multivariate setting, we propose to extend the standard location and scale mixture concept into a so called multiple scaled framework which has the advantage of allowing different tail and skewness behaviours in each dimension of the variable space with arbitrary correlation between dimensions. The approach builds upon, and develops further, previous work on scale mixtures of Gaussians [21]. Estimation of the parameters is provided via an EM algorithm with a particular focus on NIG distributions. Inference is then extended to cover the case of mixtures of such multiple scaled distributions for application to clustering. Assessments on simulated and real data confirm the gain in degrees of freedom and flexibility in modelling data of varying tail behaviour and directional shape. In addition, comparison with other similar models of GH distributions shows that the later are not as flexible as claimed.

6.2.4. Bayesian mixtures of multiple scaled distributions

Participants: Florence Forbes, Alexis Arnaud.

Joint work with: Emmanuel Barbier and Benjamin Lemasson from Grenoble Institute of Neuroscience.

In previous work [21], inference for mixtures of multiple scaled distributions has been carried out based on maximum likelihood principle and using the EM algorithm. In this work we consider a Bayesian treatment of these models for the many advantages that the Bayesian framework offers in the mixture model context. Mainly it avoids the ill-posed nature of maximum likelihood due to the presence of singularities in the likelihood function. A mixture component may collapse by becoming centered at a single data vector sending its covariance to 0 and the model likelihood to infinity. A Bayesian treatment protects the algorithm from this problem occurring in ordinary EM. Also, Bayesian model comparison embodies the principle that states that simple models should be preferred. Typically, maximum likelihood does not provide any guidance on the choice of the model order as more complex models can always fit the data better. For standard scale mixture of Gaussians, the usual Normal-Wishart prior can be used for the Gaussian parameters. For multiple scaled distributions, the specific decomposition of the covariance requires appropriate separated priors on the eigenvectors and eigenvalues of the scale matrix. Such a decomposition has been already examined in various works on priors for covariance matrix. In this work we consider several possibilities. We derive an inference scheme based on variational approximation and show how to apply this to model selection. In particular, we consider the issue of selecting automatically an appropriate number of classes in the mixtures. We show how to select this number from a single run avoiding the repetitive inference and comparison of all possible models.

6.2.5. EM for Weighted-Data Clustering

Participant: Florence Forbes.

Joint work with: Israel Gebru, Xavier Alameda-Pined and Radu Horaud from the Inria Perception team.

Data clustering has received a lot of attention and many methods, algorithms and software packages are currently available. Among these techniques, parametric finite-mixture models play a central role due to their interesting mathematical properties and to the existence of maximum-likelihood estimators based on expectation-maximization (EM). In this work we propose a new mixture model that associates a weight with each observed data point. We introduce a Gaussian mixture with weighted data and we derive two EM algorithms: the first one assigns a fixed weight to each observed datum, while the second one treats the weights as hidden variables drawn from gamma distributions. We provide a general-purpose scheme for weight initialization and we thoroughly validate the proposed algorithms by comparing them with several parametric and non-parametric clustering techniques. We demonstrate the utility of our method for clustering heterogeneous data, namely data gathered with different sensorial modalities, e.g., audio and vision. See also an application in [40].

6.3. Statistical models for Neuroscience

6.3.1. Physiologically informed Bayesian analysis of ASL fMRI data

Participants: Florence Forbes, Aina Frau Pascual, Thomas Vincent.

Joint work with: Philippe Ciuciu from Team Parietal and Neurospin, CEA in Saclay.

ASL fMRI data provides a quantitative measure of blood perfusion, that can be correlated to neuronal activation. In contrast to BOLD measure, it is a direct measure of cerebral blood flow. However, ASL data has a lower SNR and resolution so that the recovery of the perfusion response of interest suffers from the contamination by a stronger BOLD component in the ASL signal. In this work [38], [39] we consider a model of both BOLD and perfusion components within the ASL signal. A physiological link between these two components is analyzed and used for a more accurate estimation of the perfusion response function in particular in the usual ASL low SNR conditions.

6.3.2. Physiological models comparison for the analysis of ASL fMRI data

Participants: Florence Forbes, Aina Frau Pascual.

Joint work with: Philippe Ciuciu from Team Parietal and Neurospin, CEA in Saclay.

Physiological models have been proposed to describe the processes that underlie the link between neural and hemodynamic activity in the brain. Among these, the Balloon model describes the changes in blood flow, blood volume and oxygen concentration when an hemodynamic response is ensuing neural activation. Next, a *BOLD signal model* links these variables to the measured BOLD signal. Taken together, these equations allow the precise modeling of the coupling between the cerebral blood flow (CBF) and hemodynamic response (HRF). However, several competing versions of BOLD signal model have been described in the past. In this work, we compare different physiological models linking CBF to HRF and different BOLD signal models too in terms of least squares error and log-likelihood, and we assess the impact of this setting in the context of Arterial Spin Labelling (ASL) functional Magnetic Resonance Imaging (fMRI) data analysis.

6.3.3. Variational EM for the analysis of ASL fMRI data

Participants: Florence Forbes, Aina Frau Pascual.

Joint work with: Philippe Ciuciu from Team Parietal and Neurospin, CEA in Saclay.

In this work, the goal is to analyse ASL data by accounting jointly for both the BOLD and perfusion components in the signal. Using the model proposed in [77], we design a variational EM approach to estimate the model parameters as a faster alternative to the MCMC approach used in [77] and [39].

6.3.4. Metaheuristics for the analysis of fMRI data

Participants: Florence Forbes, Pablo Mesejo Santiago.

Joint work with: Jan Warnking from Grenoble Institute of Neuroscience.

The undergoing work is focused on the optimization of nonlinear models for fMRI data analysis, specially Blood-oxygen-level dependent (BOLD) MR modality. The current optimization procedure consists of a Bayesian inversion of the nonlinear model using a Gauss-Newton/Expectation-Maximization algorithm. Such an optimization procedure is time-consuming and achieves sub-optimal results. Therefore, the current research work is mainly focused on improving these results by experimenting with global search optimization methods, like metaheuristics (MHs). Secondly, MHs can also be of great help in the development of minimization algorithms for solving problems with orthogonality constraints (like in polynomial optimization, combinatorial optimization, eigenvalue problems, sparse PCA, matrix rank minimization, etc.). Thus, another main research line is concerned with the application of MHs to this problem and, if necessary, the design and implementation of new evolutionary operators that preserve orthogonality. And, finally, we are also trying to create advanced statistical models for coupling Arterial Spin Labeling (ASL) and BOLD MR modalities to study brain function.

6.3.5. Model selection for hemodynamic brain parcellation in fMRI

Participant: Florence Forbes.

Joint work with: Lotfi Chaari, Mohanad Albughdadi, Jean-Yves Tourneret from IRIT-ENSEEIHT in Toulouse and Philippe Ciuciu from Neurospin, CEA in Saclay.

Brain parcellation into a number of hemodynamically homogeneous regions (parcels) is a challenging issue in fMRI analyses. This task has been recently integrated in the joint detection-estimation (JDE) resulting in the so-called joint detection-parcellation-estimation (JPDE) model. JPDE automatically estimates the parcels from the fMRI data but requires the desired number of parcels to be fixed. This is potentially critical in that the chosen number of parcels may influence detection-estimation performance. In this paper [30], we propose a model selection procedure to automatically fix the number of parcels from the data. The selection procedure relies on the calculation of the free energy corresponding to each concurrent model, within the variational expectation maximization framework. Experiments on synthetic and real fMRI data demonstrate the ability of the proposed procedure to select an adequate number of parcels. We also investigated the use of Latent Dirichlet Processes.

6.3.6. Partial volume estimation in brain MRI revisited

Participant: Florence Forbes.

Joint work with: Alexis Roche from Siemens Advanced Clinical Imaging Technology, Department of Radiology, CHUV, Signal Processing Laboratory (LTS5), EPFL, Lausanne, Switzerland.

Image-guided diagnosis of brain disease calls for accurate morphometry algorithms, e.g., in order to detect focal atrophy patterns relating to early-stage progression of particular forms of dementia. To date, widely used brain morphometry packages rest upon discrete Markov random field (MRF) image segmentation models that ignore, or do not fully account for partial voluming, leading to potentially inaccurate estimation of tissue volumes. Although several partial volume (PV) estimation methods have been proposed in the literature from the early 90's, none of them seems to be in common use. In [43], we propose a fast algorithm to estimate brain tissue concentrations from conventional T1-weighted images based on a Bayesian maximum a posteriori formulation that extends the "mixel" model developed in the 90's. A key observation is the necessity to incorporate additional prior constraints to the "mixel" model for the estimation of plausible concentration maps. Experiments on the ADNI standardized dataset show that global and local brain atrophy measures from the proposed algorithm yield enhanced diagnosis testing value than with several widely used soft tissue labeling methods.

6.3.7. Tumor classification and prediction using robust multivariate clustering of multiparametric MRI

Participants: Florence Forbes, Alexis Arnaud.

Joint work with: Emmanuel Barbier and Benjamin Lemasson from Grenoble Institute of Neuroscience.

Advanced statistical clustering approaches are promising tools to better exploit the wealth of MRI information especially on large cohorts and multi-center studies. In neuro-oncology, the use of multiparametric MRI may better characterize brain tumor heterogeneity. To fully exploit multiparametric MRI (e.g. tumor classification), appropriate analysis methods are yet to be developed. They offer improved data quality control by allowing automatic outlier detection and improved analysis by identifying discriminative tumor signatures with measurable predictive power. In this work, we show on small animals data that advanced statistical learning approaches can help 1) in organizing existing data by detecting and excluding outliers and 2) in building a dictionary of tumor fingerprints from a clustering analysis of their microvascular features. Future work should include the integration in a joint statistical model of both automatic ROI delineation and clustering for whole brain data analysis, with a better use of anatomical information. This work has been submitted to the ISMRM 2015 conference and accepted in the SFMRMB 2015 conference [45].

6.4. Markov models

6.4.1. Identifying Interactions between Tropical Plant Species: A Correlation Analysis of High-Throughput Environmental DNA Sequence Data based on Random Matrix Theory

Participants: Florence Forbes, Angelika Studeny.

This is joint work with: Eric Coissac and Pierre Taberlet from LECA (Laboratoire d'Ecologie Alpine) and Alain Viari from Inria team Bamboo.

The study of species cooccurence pattern has always been central to community ecology. The rise of high-throughput molecular methods and their use in ecology nowadays allows for a facilitated access to new data of an unprecedented quantity. We address the question about the identification of genuine species interactions in the light of these novel data. The statistical analysis has to be tailored to the data specifics: the large amount of available data as well as biases inherent to the data extraction methods. The latter can cause spurious interactions while the former complicates any statistical modelling approach. In addition, the resolution of the data provided is rarely on the species level. In this work, we conduct a thorough correlation analysis between MOTUs (molecular operating taxonomic unit) on different spatial scales to investigate global as well as local spatial pattern. Although this type of analysis is per se exploratory, we suggest it here in order to separate true species interaction from random pattern and to identify species subgroups for further in detail modelling. A random-matrix approach allows us to derive objective cut-off values for genuine correlations. We compare the results with those derived by the application of a model-based, sparse regression approach. Our study shows that despite their seemingly less precise nature when it comes to species identification, these data enable us to reveal mechanisms that structure an ecological community. In the light of the nowadays facilitated access to molecular data, this points the way to a novel set of efficient methods for community analysis.

6.4.2. Modelling multivariate counts with graphical Markov models.

Participant: Jean-Baptiste Durand.

Joint work with: Pierre Fernique (Montpellier 2 University, CIRAD and Inria Virtual Plants) and Yann Guédon (CIRAD and Inria Virtual Plants)

Multivariate count data are defined as the number of items in different states issued from sampling within a population, which individuals own items in various numbers and states. The analysis of multivariate count data is a recurrent and crucial issue in numerous modelling problems, particularly in the fields of biology and ecology (where the data can represent, for example, children counts associated with multitype branching processes), sociology and econometrics. Denoting by K the number of states, multivariate count data analysis relies on modelling the joint distribution of the K-dimensional random vector $N = (N_0, ..., N_{K-1})$ with discrete components. Our work focused on I) Identifying states that appear simultaneously, or on the contrary that are mutually exclusive. This was achieved by identifying conditional independence relationships between the K variables; II)Building parsimonious parametric models consistent with these relationships; III) Characterizing and testing the effects of covariates on the distribution of N, particularly on the dependencies between its components.

Our context of application was characterised by zero-inflated, often right skewed marginal distributions. Thus, Gaussian and Poisson distributions were not *a priori* appropriate. Moreover, the multivariate histograms typically had many cells, most of which were empty. Consequently, nonparametric estimation was not efficient.

We developed an approach based on probabilistic graphical models (Koller & Friedman, 2009 [73]) to identify and exploit properties of conditional independence between numbers of children in different states, so as to simplify the specification of their joint distribution. The considered models are based on chain graphs. Model selection procedures are necessary to infer the graph and specify parsimonious distributions. The graph building stage was based on exploring the space of possible chain graph models, which required defining a notion of neighbourhood of these graphs. A parametric distribution was associated with each graph. It was obtained by combining families of univariate and multivariate distributions or regression models. These families were chosen by selection model procedures among different parametric families [36]. To relax the strong constraints regarding dependencies induced by using parametric distributions, mixture of graphical models were also considered [49].

Further extensions will be considered, and particularly

- Hidden Markov tree models (see 6.4.3) where the hidden state process is a multitype branching process with graphical generation distributions.
- Gaussian chain graph models, where the chain components can be identified using lasso methods.

6.4.3. Statistical characterization of tree structures based on Markov tree models and multitype branching processes, with applications to tree growth modelling.

Participant: Jean-Baptiste Durand.

Joint work with: Pierre Fernique (Montpellier 2 University and CIRAD) and Yann Guédon (CIRAD), Inria Virtual Plants.

Algorithmic issues in hidden Markov tree models were considered by Durand *et al.* (2004) [68]. This family of models was used to represent local dependencies and heterogeneity within tree-structured data. It relied on a tree-structured hidden state process, where the children states were assumed independent given their parent state. The latter assumption has been relaxed in an extension of these models and new algorithmic solutions for model inference have been proposed in Pierre Fernique's PhD [70]. An application to the study of the cell lineage in biological tissues responsible for the plant growth has been considered. In this setting, the number of children is small (between 0 and 2) and a saturated model has been considered to model transitions between parent and configurations of children states. Extensions will be proposed, based on the parametric discrete multivariate distributions developed in Section 6.4.2 .

6.4.4. Change-point models for tree-structured data

Participant: Jean-Baptiste Durand.

Joint work with: Pierre Fernique (Montpellier 2 University and CIRAD) and Yann Guédon (CIRAD), Inria Virtual Plants.

As an alternative to the hidden Markov tree models discussed in Section 6.4.3 , subtrees with similar attributes can be identified using multiple change-point models. These approaches are well-developed in the context of sequence analysis, but their extensions to tree-structured data are not straightforward. Their advantage on hidden Markov models is to relax the strong constraints regarding dependencies induced by parametric distributions and local parent-children dependencies. Heuristic approaches for change-point detection in trees were proposed and applied to the analysis of patchiness patterns (consisting of canopies made of clumps of either vegetative or flowering botanical units) in mango trees [70].

6.4.5. Hidden Markov models for the analysis of eye movements

Participant: Jean-Baptiste Durand.

Joint work with: Anne Guérin-Dugué (GIPSA-lab) and Benoit Lemaire (Laboratoire de Psychologie et Neurocognition)

In the last years, GIPSA-lab has developed computational models of information search in web-like materials, using data from both eye-tracking and electroencephalograms (EEGs). These data were obtained from experiments, in which subjects had to make some kinds of press reviews. In such tasks, reading process and decision making are closely related. Statistical analysis of such data aims at deciphering underlying dependency structures in these processes. Hidden Markov models (HMMs) have been used on eye movement series to infer phases in the reading process that can be interpreted as steps in the cognitive processes leading to decision. In HMMs, each phase is associated with a state of the Markov chain. The states are observed indirectly through eye-movements. Our approach was inspired by Simola *et al.* (2008) [76], but we used hidden semi-Markov models for better characterization of phase length distributions. The estimated HMM highlighted contrasted reading strategies (i.e., state transitions), with both individual and document-related variability.

However, the characteristics of eye movements within each phase tended to be poorly discriminated. As a result, high uncertainty in the phase changes arose, and it could be difficult to relate phases to known patterns in EEGs.

As a perspective, we aim at developing an integrated model coupling EEG and eye movements within one single HMM for better identification of the phases. Here, the coupling should incorporate some delay between the transitions in both (EEG and eye-movement) chains, since EEG patterns associated to cognitive processes occur lately with respect to eye-movement phases. Moreover, EEGs and scanpaths were recorded with different time resolutions, so that some resampling scheme must be added into the model, for the sake of synchronizing both processes. Probabilistic graphical models (see Section 6.4.2) will be inferred from the channel correlations to represent interactions between brain zones. The variability of these graphs is partly explained by individual differences in text exploration, which will have to be quantified.

6.4.6. Hyper-Spectral Image Analysis with Partially-Latent Regression and Spatial Markov Dependencies

Participant: Florence Forbes.

Joint work with: Antoine Deleforge, Sileye Ba and Radu Horaud from the Inria Perception team.

Hyper-spectral data can be analyzed to recover physical properties at large planetary scales. This involves resolving inverse problems which can be addressed within machine learning, with the advantage that, once a relationship between physical parameters and spectra has been established in a data-driven fashion, the learned relationship can be used to estimate physical parameters for new hyper-spectral observations. Within this framework, we propose a spatially-constrained and partially-latent regression method which maps high-dimensional inputs (hyper-spectral images) onto low-dimensional responses (physical parameters). The proposed regression model comprises two key features. Firstly, it combines a Gaussian mixture of locally-linear mappings (GLLiM) with a partially-latent response model described in [17]. While the former makes high-dimensional regression tractable, the latter enables to deal with physical parameters that cannot be observed or, more generally, with data contaminated by experimental artifacts that cannot be explained with noise models. Secondly, spatial constraints are introduced in the model through a Markov random field (MRF) prior which provides a spatial structure to the Gaussian-mixture hidden variables. Experiments conducted on a database composed of remotely sensed observations collected from the Mars planet by the Mars Express orbiter demonstrate the effectiveness of the proposed model. A preliminary version of the work can be found in [31].

6.5. Semi and non-parametric methods

6.5.1. Conditional extremal events

Participant: Stéphane Girard.

Joint work with: L. Gardes (Univ. Strasbourg), A. Daouia (Univ. Toulouse I and Univ. Catholique de Louvain), J. Elmethni (Univ. Paris 5) and S. Louhichi (Univ. Grenoble 1)

The goal of the PhD thesis of Alexandre Lekina was to contribute to the development of theoretical and algorithmic models to tackle conditional extreme value analysis, *ie* the situation where some covariate information X is recorded simultaneously with a quantity of interest Y. In such a case, the tail heaviness of Y depends on X, and thus the tail index as well as the extreme quantiles are also functions of the covariate. We combine nonparametric smoothing techniques [71] with extreme-value methods in order to obtain efficient estimators of the conditional tail index and conditional extreme quantiles. The strong consistency of such estimator is established in [53]. When the covariate is functional and random (random design) we focus on kernel methods [58].

Conditional extremes are studied in climatology where one is interested in how climate change over years might affect extreme temperatures or rainfalls. In this case, the covariate is univariate (time). Bivariate examples include the study of extreme rainfalls as a function of the geographical location. The application part of the study is joint work with the LTHE (Laboratoire d'étude des Transferts en Hydrologie et Environnement) located in Grenoble.

6.5.2. Estimation of extreme risk measures

Participant: Stéphane Girard.

Joint work with: E. Deme (Univ. Gaston-Berger, Sénégal, J. Elmethni (Univ. Paris 5), L. Gardes and A. Guillou (Univ. Strasbourg)

One of the most popular risk measures is the Value-at-Risk (VaR) introduced in the 1990's. In statistical terms, the VaR at level $\alpha \in (0,1)$ corresponds to the upper α-quantile of the loss distribution. The Value-at-Risk however suffers from several weaknesses. First, it provides us only with a pointwise information: VaR(α) does not take into consideration what the loss will be beyond this quantile. Second, random loss variables with light-tailed distributions or heavy-tailed distributions may have the same Value-at-Risk . Finally, Value-at-Risk is not a coherent risk measure since it is not subadditive in general. A coherent alternative risk measure is the Conditional Tail Expectation (CTE), also known as Tail-Value-at-Risk, Tail Conditional Expectation or Expected Shortfall in case of a continuous loss distribution. The CTE is defined as the expected loss given that the loss lies above the upper α-quantile of the loss distribution. This risk measure thus takes into account the whole information contained in the upper tail of the distribution. It is frequently encountered in financial investment or in the insurance industry. In [52], we have established the asymptotic properties of the CTE estimator in case of extreme losses, *i.e.* when $\alpha \to 0$ as the sample size increases. We have exhibited the asymptotic bias of this estimator, and proposed a bias correction based on extreme-value techniques. In [20], we study the situation where some covariate information is available. We thus has to deal with conditional extremes (see paragraph 6.5.1). We also proposed a new risk measure (called the Conditional Tail Moment) which encompasses various risk measures, such as the CTE, as particular cases.

6.5.3. Multivariate extremal events

Participants: Stéphane Girard, Gildas Mazo, Florence Forbes.

Joint work with: C. Amblard (TimB in TIMC laboratory, Univ. Grenoble I), L. Gardes (Univ. Strasbourg) and L. Menneteau (Univ. Montpellier II)

Copulas are a useful tool to model multivariate distributions [75]. At first, we developed an extension of some particular copulas [1]. It followed a new class of bivariate copulas defined on matrices [55] and some analogies have been shown between matrix and copula properties.

However, while there exist various families of bivariate copulas, much fewer has been done when the dimension is higher. To this aim an interesting class of copulas based on products of transformed copulas has been proposed in the literature. The use of this class for practical high dimensional problems remains challenging. Constraints on the parameters and the product form render inference, and in particular the likelihood computation, difficult. We proposed a new class of high dimensional copulas based on a product of transformed bivariate copulas [64]. No constraints on the parameters refrain the applicability of the proposed class which is well suited for applications in high dimension. Furthermore the analytic forms of the copulas within this class allow to associate a natural graphical structure which helps to visualize the dependencies and to compute the likelihood efficiently even in high dimension. The extreme properties of the copulas are also derived and an R package has been developed.

As an alternative, we also proposed a new class of copulas constructed by introducing a latent factor. Conditional independence with respect to this factor and the use of a nonparametric class of bivariate copulas lead to interesting properties like explicitness, flexibility and parsimony. In particular, various tail behaviours are exhibited, making possible the modeling of various extreme situations [42]. A pairwise moment-based inference procedure has also been proposed and the asymptotic normality of the corresponding estimator has been established [66].

In collaboration with L. Gardes, we investigate the estimation of the tail copula which is widely used to describe the amount of extremal dependence of a multivariate distribution. In some situations such as risk management, the dependence structure can be linked with some covariate. The tail copula thus depends on this covariate and is referred to as the conditional tail copula. The aim of our work is to propose a nonparametric estimator of the conditional tail copula and to establish its asymptotic normality [57].

6.5.4. Level sets estimation

Participant: Stéphane Girard.

Joint work with: A. Guillou and L. Gardes (Univ. Strasbourg), A. Nazin (Univ. Moscou), G. Stupfler (Univ. Aix-Marseille) and A. Daouia (Univ. Toulouse I and Univ. Catholique de Louvain)

The boundary bounding the set of points is viewed as the larger level set of the points distribution. This is then an extreme quantile curve estimation problem. We proposed estimators based on projection as well as on kernel regression methods applied on the extreme values set, for particular set of points [10]. We also investigate the asymptotic properties of existing estimators when used in extreme situations. For instance, we have established in collaboration with G. Stupfler that the so-called geometric quantiles have very counter-intuitive properties in such situations [63], [62] and thus should not be used to detect outliers. These resuls are submitted for publication.
In collaboration with A. Daouia, we investigate the application of such methods in econometrics [16]: A new characterization of partial boundaries of a free disposal multivariate support is introduced by making use of large quantiles of a simple transformation of the underlying multivariate distribution. Pointwise empirical and smoothed estimators of the full and partial support curves are built as extreme sample and smoothed quantiles. The extreme-value theory holds then automatically for the empirical frontiers and we show that some fundamental properties of extreme order statistics carry over to Nadaraya's estimates of upper quantile-based frontiers.
In collaboration with A. Nazin, we define new estimators of the frontier function based on linear programming methods. The frontier is defined as the solution of a linear optimization problem under inequality constraints. The estimator is shown to be strongly consistent with respect to the L_1 norm and we establish that it reaches the optimal minimax rate of convergence [58].

In collaboration with G. Stupfler and A. Guillou, new estimators of the boundary are introduced. The regression is performed on the whole set of points, the selection of the "highest" points being automatically performed by the introduction of high order moments [22].

6.5.5. Retrieval of Mars surface physical properties from OMEGA hyperspectral images.

Participants: Stéphane Girard, Alessandro Chiancone.

Joint work with: S. Douté from Laboratoire de Planétologie de Grenoble, J. Chanussot (Gipsa-lab and Grenoble-INP) and J. Saracco (Univ. Bordeaux).

Visible and near infrared imaging spectroscopy is one of the key techniques to detect, to map and to characterize mineral and volatile (eg. water-ice) species existing at the surface of planets. Indeed the chemical composition, granularity, texture, physical state, etc. of the materials determine the existence and morphology of the absorption bands. The resulting spectra contain therefore very useful information. Current imaging spectrometers provide data organized as three dimensional hyperspectral images: two spatial dimensions and one spectral dimension. Our goal is to estimate the functional relationship F between some observed spectra and some physical parameters. To this end, a database of synthetic spectra is generated by a physical radiative transfer model and used to estimate F. The high dimension of spectra is reduced by Gaussian regularized sliced inverse regression (GRSIR) to overcome the curse of dimensionality and consequently the sensitivity of the inversion to noise (ill-conditioned problems) [15]. We have also defined an adaptive version of the method which is able to deal with block-wise evolving data streams [13].
In his PhD thesis work, Alessandro Chiancone studies the extension of the SIR method to different sub-populations. The idea is to assume that the dimension reduction subspace may not be the same for different clusters of the data [46]. He also published a paper on a previous work in the field of hierarchical segmentation of images [14].

MNEMOSYNE Project-Team

6. New Results

6.1. Overview

Though our view is systemic, our daily research activities are concerned with the design, at a given scale of description, of models of neuronal structures, each concerned with a specific learning paradigm. Of course, a major challenge is to keep in mind the systemic view, to put a specific emphasis on the way each neuronal structure communicates with the rest of the system and to highlight how the learning paradigm interplays with other memory systems.

Among the numerous loops involving the brain, the body and the environment, a basic grid of description corresponds to distinguish “Perception Loops”, the goal of which is to extract from the inner and outer world sensory invariants helpful to identify and evaluate the current state and to make predictions from previous learning, and “Action Loops”, the goal of which is to rely on this sensory, emotional and motivational information to decide, plan and trigger actions for the benefit of the body.

Presently, our team is engaged on the following topics: Concerning perception loops, we are firstly considering the role of the hippocampus and of the posterior cortex in learning high level sensory cues that contribute to pavlovian conditioning in the amygdala. Secondly, we are investigating the role of the thalamus in attentional shifts in the cortex. Concerning Action loops, we are preparing a critical analysis of the current views of the interactions between the prefrontal cortex and the basal ganglia. Finally, we also report here more methodological achievements.

6.2. Pavlovian conditioning

Pavlovian conditioning is an outstanding example of a systemic process involving several cerebral structures and several modes of learning. This year, we have made more precise our model of amygdala [7], with a special emphasis on the variety of its inputs, including by neuromodulation [12]. We have also specifically discussed and contrasted the role of cerebral structures involved in this learning, from the point of view of information processing [11]. In addition to the amygdala, the structures of interest are the hippocampus and the posterior and prefrontal cortex and begin to be investigated in ongoing studies.

6.3. The thalamus is more than a relay

Many recent results in neuroscience indicate that the role of the thalamus in the brain is certainly more important than it used to be considered, particularly concerning its relation with the cortex. Our modeling and bibliographic studies were carried out in the Keops project (*cf.* § 7.2) with our chilean neuroscientist colleagues studying non standard ganglion cells in the retina. Particularly, the PhD work by Carlos Carvajal [1] led us to propose a biologically-founded algorithm for the interplay between the modulatory and driving connections between the thalamus and the cortex, with the strong constraint of proposing a system working on a real visual flow.

6.4. On the computational efficiency of Basal Ganglia models

Many valuable models have been proposed to capture the richness of the fundamental relations between the basal ganglia and prominent brain structures including the prefrontal cortex, the hippocampus and the superior colliculus. To choose among them the mechanisms on which to build the design of the motor pole of our brain-inspired system, a fundamental issue is to evaluate the efficiency of these models in more realistic cases than the ones which are generally considered by the authors [24]. For this reason, we have conducted a systematic study of several basal ganglia computational models to check of their scalability in terms of ation representational space [25]. Unfortunately, we found most of them to not be scalable and some of them to not be reproducible at all.

Another way to explore the computational efficiency of neuronal models is to implement them at lower levels of description. This is currently being done with one model developed in our lab at a level corresponding to a neuronal assembly with a mean activity expressed using a single variable. This mesoscopic approach has been refined to a microscopic scale description level, i.e taking into account individual neurons and synapses. Besides the confirmation of many of the results of initial model with a more detailed formalism, this new model has allowed us to highlight the facilitating role of inhibitory interneurons in the decision-making and action selection processes.

6.5. Distributed Self-Organization

The formation of the sensory homunculus in the primary sensory cortex (SI) is believed to be the result of a dynamic neural self-organization process that starts before birth and lasts for several years, allowing the brain to cope with sensory or brain lesions. The exact neural mechanisms driving this self-organization are not yet known and the role of the somatosensory attention remains unclear in this picture. We thus investigated the influence of somatosensory attention onto the two-dimensional structure of area 3b neuronal receptive fields (RFs) using a computational model [2] based on the dynamic neural field theory. This computational model of SI (area 3b) is able to explain experimental data in the monkey and hypothesizes role for the somatosensory attention in the shaping of SI receptive fields.

MOAIS Project-Team

5. New Results

5.1. Scheduling semi-malleable jobs to minimize mean flow time

This paper [9] deals with the problem of scheduling n_A malleable and n_B non-malleable jobs to be executed together on two parallel identical machines to minimize mean flow time. We propose a set of dominant schedules for this problem, and a dynamic programming algorithm that finds an optimal schedule in this dominant set in time $O(n_A^2 n_B)$.

5.2. Elements of Design for Containers and Solutions in the LinBox Library

We describe in this paper [12] new design techniques used in the exact linear algebra library LinBox, intended to make the library safer and easier to use, while keeping it generic and efficient. First, we review the new simplified structure for containers, based on our *founding scope allocation* model. We explain design choices and their impact on coding: unification of our matrix classes, clearer model for matrices and submatrices,... Then we present a variation of the *strategy* design pattern that is comprised of a controller-plugin system: the controller (solution) chooses among plug-ins (algorithms) that always call back the controllers for subtasks. We give examples using the solution `mul`. Finally we present a benchmark architecture that serves two purposes: Providing the user with easier ways to produce graphs; Creating a framework for automatically tuning the library and supporting regression testing.

5.3. Scheduling Data Flow Program in XKaapi: A New Affinity Based Algorithm for Heterogeneous Architectures

Efficient implementations of parallel applications on heterogeneous hybrid architectures require a careful balance between computations and communications with accelerator devices. Even if most of the communication time can be overlapped by computations, it is essential to reduce the total volume of communicated data. The literature therefore abounds with ad hoc methods to reach that balance, but these are architecture and application dependent. We propose [12] here a generic mechanism to automatically optimize the scheduling between CPUs and GPUs, and compare two strategies within this mechanism: the classical Heterogeneous Earliest Finish Time (HEFT) algorithm and our new, parametrized, Distributed Affinity Dual Approximation algorithm (DADA), which consists in grouping the tasks by affinity before running a fast dual approximation. We ran experiments on a heterogeneous parallel machine with twelve CPU cores and eight NVIDIA Fermi GPUs. Three standard dense linear algebra kernels from the PLASMA library have been ported on top of the XKaapi runtime system. We report their performances. It results that HEFT and DADA perform well for various experimental conditions, but that DADA performs better for larger systems and number of GPUs, and, in most cases, generates much lower data transfers than HEFT to achieve the same performance.

5.4. Evaluation of OpenMP Dependent Tasks with the KASTORS Benchmark Suite

The recent introduction of task dependencies in the OpenMP specification provides new ways of synchronizing tasks. Application programmers can now describe the data a task will read as input and write as output, letting the runtime system resolve fine-grain dependencies between tasks to decide which task should execute next. Such an approach should scale better than the excessive global synchronization found in most OpenMP 3.0 applications. As promising as it looks however, any new feature needs proper evaluation to encourage application programmers to embrace it. This paper [26] introduces the KASTORS benchmark suite designed to evaluate OpenMP tasks dependencies. We modified state-of-the-art OpenMP 3.0 benchmarks and data-flow parallel linear algebra kernels to make use of tasks dependencies. Learning from this experience, we propose extensions to the current OpenMP specification to improve the expressiveness of dependencies. We eventually evaluate both the GCC/libGOMP and the CLANG/libIOMP implementations of OpenMP 4.0 on our KASTORS suite, demonstrating the interest of task dependencies compared to taskwait-based approaches.

5.5. Sparse Polynomial Interpolation Codes and their decoding beyond half the minimal distance

We present [21] algorithms performing sparse univariate polynomial interpolation with errors in the evaluations of the polynomial. Based on the initial work by Comer, Kaltofen and Pernet [Proc. ISSAC 2012], we define the sparse polynomial interpolation codes and state that their minimal distance is precisely the length divided by twice the sparsity. At ISSAC 2012, we have given a decoding algorithm for as much as half the minimal distance and a list decoding algorithm up to the minimal distance. Our new polynomial-time list decoding algorithm uses sub-sequences of the received evaluations indexed by a linear progression, allowing the decoding for a larger radius, that is, more errors in the evaluations while returning a list of candidate sparse polynomials. We quantify this improvement for all typically small values of number of terms and number of errors, and provide a worst case asymptotic analysis of this improvement. For instance, for sparsity T = 5 with up to 10 errors we can list decode in polynomial-time from 74 values of the polynomial with unknown terms, whereas our earlier algorithm required 2T (E + 1) = 110 evaluations. We then propose two variations of these codes in characteristic zero, where appropriate choices of values for the variable yield a much larger minimal distance: the length minus twice the sparsity.

5.6. A Spatiotemporal Data Aggregation Technique for Performance Analysis of Large-scale Execution Traces

Analysts commonly use execution traces collected at runtime to understand the behavior of an application running on distributed and parallel systems. These traces are inspected post mortem using various visualization techniques that, however, do not scale properly for a large number of events. This issue, mainly due to human perception limitations, is also the result of bounded screen resolutions preventing the proper drawing of many graphical objects. This paper [21] proposes a new visualization technique overcoming such limitations by providing a concise overview of the trace behavior as the result of a spatiotemporal data aggregation process. The experimental results show that this approach can help the quick and accurate detection of anomalies in traces containing up to two hundred million events.

5.7. Scheduling independent tasks on multi-cores with GPU accelerators

More and more computers use hybrid architectures combining multi-core processors and hardware accelerators like GPUs (Graphics Processing Units). We present in this paper [3] a new method for scheduling efficiently parallel applications with m CPUs and k GPUs, where each task of the application can be processed either on a core (CPU) or on a GPU. The objective is to minimize the maximum completion time (makespan). The corresponding scheduling problem is NP-hard, we propose an efficient approximation algorithm which achieves an approximation ratio of $4/3 + 1/3k$. We first detail and analyze the method, based on a dual approximation scheme, that uses dynamic programming to balance evenly the load between the heterogeneous resources. Then, we present a faster approximation algorithm for a special case of the previous problem, where all the tasks are accelerated when affected to GPU, with a performance guarantee of $3/2$ for any number of GPUs. We run some simulations based on realistic benchmarks and compare the solutions obtained by a relaxed version of the generic method to the one provided by a classical scheduling algorithm (HEFT). Finally, we present an implementation of the $4/3$-approximation and its relaxed version on a classical linear algebra kernel into the scheduler of the XKaapi runtime system.

5.8. A Flexible Framework for Asynchronous In Situ and In Transit Analytics for Scientific Simulations

High performance computing systems are today composed of tens of thousands of processors and deep memory hierarchies. The next generation of machines will further increase the unbalance between I/O capabilities and processing power. To reduce the pressure on I/Os, the in situ analytics paradigm proposes to process the data as closely as possible to where and when the data are produced. Processing can be embedded in the simulation code, executed asynchronously on helper cores on the same nodes, or performed in transit on staging nodes dedicated to analytics. Today, software environnements as well as usage scenarios still need to be investigated before in situ analytics become a standard practice. In this paper [3] we introduce a framework for designing, deploying and executing in situ scenarios. Based on a component model, the scientist designs analytics workflows by first developing processing components that are next assembled in a dataflow graph through a Python script. At runtime the graph is instantiated according to the execution context, the framework taking care of deploying the application on the target architecture and coordinating the analytics workflows with the simulation execution. Component coordination, zero-copy intra-node communications or inter-nodes data transfers rely on per-node distributed daemons. We evaluate various scenarios performing in situ and in transit analytics on large molecular dynamics systems simulated with Gromacs using up to 1664 cores. We show in particular that analytics processing can be performed on the fraction of resources the simulation does not use well, resulting in a limited impact on the simulation performance (less than 6%). Our more advanced scenario combines in situ and in transit processing to compute a molecular surface based on the Quicksurf algorithm.

5.9. Generic Deterministic Random Number Generation in Dynamic-Multithreaded Platforms

On dynamic multithreaded platforms with on-line scheduling such as work-stealing, randomized computations raise the issue of reproducibility. Compliant with de facto standard sequential Deterministic Random Number Generators (DRNGs) noted R, we propose [23] a parallel DRNG implementation for finite computations that provides deterministic parallel execution. It uses the stateless sub-stream approach, enabling the use of efficient DRNG such as Mersenne Twister or Linear Congruential. We demonstrate that if R provides fast jump ahead in the random sequence, the re-seeding overhead is small, polylog in expectation, independently from the parallel computation's depth. Experiments benchmark the performance of randomized algorithms employing our solution against the stateful DRNG DotMix, tailored to the Cilk Plus dynamic multithreading runtime. The overhead of our implementation ParDRNG compares favorably to the linear overhead of DotMix re-seedings.

MODAL Project-Team

6. New Results

6.1. Highlights of the Year

Thanks to the development technological action MPAGenomics, the team has created one of the first french instances of Galaxy publicly available on the French Bioinformatics cloud. This instance is original as it offers complex statistical tools for genomic data analysis in a user-friendly interface (see 5.9).
The team obtained bilateral contracts with companies as Auchan or RougeGorge thanks to its just emerging, but promising, clustering software MixtComp (see 5.14), dedicated to full mixed and missing data.

6.2. Model for conditionally correlated categorical data

Participants: Christophe Biernacki, Matthieu Marbac-Lourdelle, Vincent Vandewalle.

An extension of the latent class model is proposed for clustering categorical data by relaxing the classical class conditional independence assumption of variables. In this model (called CCM for Conditional Correlated Model), variables are grouped into inter-independent and intra-dependent blocks in order to consider the main intra-class correlations. The dependence between variables grouped into the same block is taken into account by mixing two extreme distributions, which are respectively the independence and the maximum dependence ones. In the conditionally correlated data case, this approach is expected to reduce biases involved by the latent class model and to produce a meaningful model with few additional parameters. The parameters estimation by maximum likelihood is performed by an EM algorithm while a MCMC algorithm avoiding combinatorial problems involved by the block structure search is used for model selection. Applications on sociological and biological data sets bring out the proposed model interest. These results strengthen the idea that the proposed model is meaningful and that biases induced by the conditional independence assumption of the latent class model are reduced. This work has been now accepted in an international journal [24]. Furthermore, an R package (Clustericat) is available on CRAN (see 5.3).

6.3. Model for conditionally correlated categorical data

Participants: Christophe Biernacki, Matthieu Marbac-Lourdelle, Vincent Vandewalle.

It is a model-based clustering proposal (called CMM for Conditional Modes Model) where categorical data are grouped into conditionally independent blocks. The corresponding block distribution is a parsimonious multinomial distribution where the few free parameters correspond to the most likely modality crossings, while the remaining probability mass is uniformly spread over the other modality crossings. The exact computation of the integrated complete-data likelihood allows to perform the model selection, by a Gibbs sampler, reducing the computing time consuming by parameter estimation and avoiding BIC criterion biases pointed out by our experiments. An article has been now submitted to an international journal [49]. Furthermore, an R package (CoModes) is available on Rforge (see 5.4).

6.4. Mixture model for mixed kind of data

Participants: Christophe Biernacki, Matthieu Marbac-Lourdelle, Vincent Vandewalle.

A mixture model of Gaussian copula allows to cluster mixed kind of data. Each component is composed by classical margins while the conditional dependencies between the variables is modeled by a Gaussian copula. The parameter estimation is performed by a Gibbs sampler. An article has been presented to an international conference [48] and has been also submitted to an international journal [50]. Furthermore, an R package (MixCluster) is available on Rforge (see 5.12).

6.5. Mixture of Gaussians with Missing Data

Participants: Christophe Biernacki, Vincent Vandewalle.

The generative models allow to handle missing data. This can be easily performed by using the EM algorithm, which has a closed form M-step in the Gaussian setting. This can for instance be useful for distance estimation with missing data. It has been proposed to improve the distance estimation by fitting a mixture of Gaussian distributions instead of a considering only one Gaussian component [16]. This is a joint work with Emil Eirola and Amaury Lendrasse .

6.6. Clustering and variable selection in regression

Participants: Christophe Biernacki, Loïc Yengo, Julien Jacques.

A new framework is proposed to address the issue of simultaneous linear regression and clustering of predictors where regression coefficients are assumed to be drawn from a Gaussian mixture distribution. Prediction is thus performed using the conditional distribution of the regression coefficients given the data, while clusters are easily derived from posterior distribution in groups given the data. This work is now published in [27]. Furthermore, an R package (clere) is available on Rforge (see 5.2) and an improved version of the initial model has been submitted to an international journal [52].

6.7. Model-based clustering for multivariate partial ranking data

Participants: Christophe Biernacki, Julien Jacques.

The first model-based clustering algorithm dedicated to multivariate partial ranking data is now published in an internation journal [19]. This is an extension of the (ISR) model for ranking data published in 2013. The proposed algorithm has allowed to exhibit regional alliances between European countries in the Eurovision contest, which are often suspected but never proved.

6.8. Generative models for correlated variables in regression

Participants: Christophe Biernacki, Clément Théry.

Linear regression outcomes (estimates, prevision) are known to be damaged by highly correlated covariates. However most modern datasets are expected to mechanically convey more and more highly correlated covariates due to the global increase of the amount of variables they contain. We propose to explicitly model such correlations by a family of linear regressions between the covariates. It leads to a particular generative model through the distribution explicitly introduced between correlated covariates. It has been presented to a conference [32] and is currently written as a research paper [51]. Furthermore, an R package (CorReg) is available on CRAN (see 5.5). Extension is now available for missing covariables also. It is a joint work with Gaétan Loridant.

6.9. Model-based clustering for multivariate partial ordinal data

Participants: Christophe Biernacki, Julien Jacques.

We design the first univariate probability distribution for ordinal data which strictly respects the ordinal nature of data. More precisely, it relies only on order comparisons between modalities, the proposed distribution being obtained by modeling the data generating process which is assumed, from optimality arguments, to be a stochastic binary search algorithm in a sorted table. The resulting distribution is natively governed by two meaningful parameters (position and precision) and has very appealing properties: decrease around the mode, shape tuning from uniformity to a Dirac, identifiability. Moreover, it is easily estimated by an EM algorithm since the path in the stochastic binary search algorithm is missing. Using then the classical latent class assumption, the previous univariate ordinal model is straightforwardly extended to model-based clustering for multivariate ordinal data. Again, parameters of this mixture model are estimated by an EM algorithm. Both simulated and real data sets illustrate the great potential of this model by its ability to parsimoniously identify particularly relevant clusters which were unsuspected by some traditional competitors. This work is currently in revision in an international journal [38].

6.10. Clustering for functional data into discriminative subspaces

Participant: Julien Jacques.

This is a joint work with Charles Bouveyron (Paris 5) and Etienne Côme (Inrets).

A model-based clustering method for time series has been developed, based on a discriminative functional mixture model which allows the clustering of the data in a functional subspace. This model presents the advantage to be parsimonious and can therefore handle long time series. This model has been used for analyzing different bike sharing systems In Europe.

6.11. Degeneracy in multivariate Gaussian mixtures

Participant: Christophe Biernacki.

In the case of Gaussian mixtures, unbounded likelihood is an important theoretical and practical problem. Using the weak information that the latent sample size of each component has to be greater than the space dimension, we derive a simple non-asymptotic stochastic lower bound on variances. We prove also that maximizing the likelihood under this data-driven constraint leads to consistent estimates. This work has been presented to a conference [31]. This is a joint work with Gwënaelle Castellan.

6.12. Auto-Associative Models

Participant: Serge Iovleff.

Auto-Associative models cover a large class of methods used in data analysis, among them are for example the famous PCA and the auto-associative neural networks. We describe the general properties of these models when the projection component is linear and we propose and test an easy to implement Probabilistic Semi-Linear Auto-Associative model in a Gaussian setting. This work is now published in [18].

6.13. Resampling and density estimation

Participant: Alain Celisse.

We characterized the behavior of cross-validation (Lpo) in density estimation with the L^2-loss. We considered two aspects: risk estimation and model selection. For the first one, we settled leave-one-out is optimal. On the contrary for the second one, we provided the first guidelines toward an optimal choice of the parameter p. In particular, this choice depends on the convergence rate of the best estimator in the family we consider.

6.14. Resampling and classification

Participant: Alain Celisse.

This is a joint work with Tristan Mary-Huard (INRA).

We extended known results about leave-one-out to the case of leave-p-out for the k-nearest neighbor estimator in classification with the 0-1 loss. In particular, our strategy relies on the relationship between leave-p-out and U-statistics. We derive upper bounds on the moments on the leave-p-out estimator as well as an exponential concentration inequality.

6.15. Kernel change-point

Participants: Alain Celisse, Guillemette Marot.

This is a joint work with Guillem Rigaill and Morgane Pierre-Jean (Univ. Evry).

Based on a previous work, we successfully applied kernel methods to change-point detection in the context of next generation sequencing with multivariate complex data. We also provided greatly improved algorithm in terms of computational complexity (both in time and space). With very huge amounts of data, we also suggest a new strategy based on the idea of approximating the Gram matrix by a low-rank matrix, which leads to a linear time complexity algorithm.

6.16. Normality test in RKHS

Participants: Alain Celisse, Jérémie Kellner.

In the kernel method framework, we use the MMD (maximum mean discrepancy) to derive a new goodness-of-fit test that can be used in the RKHS. When applied to the usual R^d setting, our test does not seem too sensitive to any increase on the dimension d unlike other ongoing approaches. With an infinite dimension RKHS, it exhibits a good power for a prescribed level of type-I error control.

6.17. Differential meta-analysis of RNA-seq data from multiple studies

Participant: Guillemette Marot.

This is a joint work with Andrea Rau and Florence Jaffrézic (INRA, Jouy-en-Josas).

An adaptation of meta-analysis methods initially proposed for microarray studies has been proposed for RNA-seq data. The research paper has been published in [26] and the associated R package metaRNASeq is now available on CRAN (see 5.11).

6.18. Multi-patient analysis of genomic markers

Participants: Quentin Grimonprez, Samuel Blanck, Guillemette Marot, Alain Celisse.

Tests performed during Development Technological Action MPAgenomics have shown on real data that it was also important to suggest automatic and appropriate calibrations for parameters in segmentation methods than to look for common markers able to predict patient's response. In the R package MPAgenomics (see 5.15), we have thus proposed two independent pipelines described in [17]. The choice of a given pipeline depends on the heterogeneity degree of studied genomic profiles.

6.19. Scan statistics for dependent data

Participants: Alexandru Amarioarei, Cristian Preda.

Dependent models of type block factors are introduced for scan statistics as an extension of the models based on the independent and identically distributed assumption. Approximations and errors are derived for one and two dimensions. Matlab software has been developed for this purpose.

MODEMIC Project-Team

6. New Results

6.1. Highlights of the Year

Yeasts play a central role in the wine making process. To study the yeasts in a stable environment and physiological state, a Multi-Stage Continuous Fermentor (MSCF) has been designed by the research Unit SPO (Sciences For Oenology). This device mimics the steps of the batch fermentation process. In this paper, the problem of the control of the sugar concentrations in each of the four reactors of the MSCF is considered. The cascade structure of the device leads to a constraint on the input flow rates (the control variables). A control strategy based on a linearizing control law coupled with a state observer and an anti windup component is proposed and finally implemented on the experimental process (see also 6.3.2).

BEST PAPER AWARD :

[41] **19th IFAC World Congress 2014**. C. CASENAVE, D. DOCHAIN, J. HARMAND, M. PEREZ, A. RAPAPORT, J.-M. SABLAYROLLES.

6.2. Mathematical models for microbial ecology

6.2.1. Differential equations models

Participants: Céline Casenave, Jérôme Harmand, Claude Lobry, Alain Rapaport, Alejandro Maximiliano Rojas.

Anaerobic digestion refers to the transformation of biodegradable material by micro-organisms in absence of oxygen (it can be found in waste-water treatments or industrial fermentation, and occurs naturally in soils). It receives an increasing consideration due to recent technological advances, but also because it is a source of renewable energy (bio-gas, fuel...). The anaerobic digestion is a complex set of bio-processes, for which there is a strong expectation of tractable models. We have proposed and studied new mathematical models that takes into account the following features:

- The available anaerobic digestion models used for control purposes do usually only consider soluble matter. In fact, part of the pollutants are not soluble but are under a particulate form. In order to establish whether adding the dynamics of such matter into the models is important for the system behavior or not, we have studied new models and established that depending on the kinetics of this additional reaction step, the qualitative behavior of the process may be significantly modified [26].
- Microbial food chains are present in anaerobic digestion where the different reaction steps can be seen as such: the waste products of the organisms at one trophic level (i.e. one reaction step) are consumed by organisms at the next trophic level (i.e. the next reaction step). In [55] we study a model of a two-tiered microbial 'food chain' with feedback inhibition, which was recently presented [63] as a reduced and simplified version of the anaerobic digestion model ADM1 of the International Water Association (IWA). It is known that in the absence of maintenance (or decay) the microbial 'food chain' is stable. In [63], using a purely numerical approach and ADM1 consensus parameter values, it was shown that the model remains stable when decay terms are added. In [55] we prove that introducing decay in the model preserves stability whatever its parameters values are and for a wide range of kinetics.

For the study of spatial heterogeneity in the models, we have carried on mathematical analyses of the properties of interconnected chemostats, in particular when growth rates present a substrate inhibition. In addition to the stabilizability properties discovered last year on "buffered" interconnections [33], we have studied this year yielding performances at steady state for the stabilizing configurations and characterized the set of the most efficient ones. For such configurations, we have shown that under certain circumstances, a "by-pass" of the main tank could be the best solution to ensure a global stabilization.

We have also analyzed two kinds of models, suited to specific characteristics of the microbial activity in soils:

- In [30], we have studied analytically and numerically a piece-wise linear model of carbon mineralization by two functional groups of micro-organisms in view of predicting the "priming effect" in soil ecosystems. The conclusion is that under a climate change, the augmentation of $C0^2$ will not affect to primary production and carbon storage when the plants are limited by nitrogen, but surprisingly a higher carbon input in soil should lead to a deplete of sequestered carbon and the increase of nitrogen release.
- In collaboration with Géosciences Rennes (Jean-Raynald de Dreuzy, Tristan Babey) and in the scope of the co-supervision of the PhD of Alejandro Rojas (also in the collaboration within the associated team with Chile), we investigate the equivalence between networks that represent interconnections of mobile/immobile zones in mass transfer models for soil ecosystems. For Structured INteracting Continua (SINC) models, that are described as the combination of a finite number of diffusion-dominated interconnected immobile zones exchanging with an advection-dominated mobile domain, we have proved an equivalence with Multi-Rate Mass Transfer (MRMT) and proposed a method for the identification of the equivalent MRMT model [14]. Moreover, we have shown the role of the controllability properties of a sub-system, in addition to the irreducibility of the network graph, for the input/output equivalence between several representations (work in preparation).

6.2.2. Stochastic and hybrid discrete-continuous dynamical models

Participants: Fabien Campillo, Bertrand Cloez, Coralie Fritsch.

Hybrid mass-structured chemostat models

Within the context of Coralie Fritsch thesis [12], we adopt a new modeling approach where instead on focusing on one type of model we propose different models and their interconnections, on the numerical viewpoint as well as the analytical one. Namely we propose an hybrid model of the chemostat where the population of bacteria is individually-based, each individual being described by its mass, and the subtract concentration is represented as a classic differential equation. We proved the convergence of this model in high population size toward an integro-differential system [20]. We proposed specific numerical schemes for the two approaches (see 5.3) [27].

Evolutionary invasion analysis and simulation for the chemostat

Still in the context of Coralie Fritsch thesis [12] and following her last year stay at the University of Helsinki in Otso Ovaskainen's Research Group, we consider an hybrid mass-structured mass-structured chemostat models with trait. The trait could for example be the factor of mass dissymmetry in the binary fission of a bacteria. In this context we proved an equivalence between invasion fitness for the hybrid-IBM model and the integro-differential system. We also numerically exhibit an evolutionarily singular strategy: with this given trait a monomorphic resident population cannot be invaded by a mutated population; the result is true for the two models.

PDE and stochastic models

In collaboration with M. Joannides and I. Larramendy-Valverde (I3M, University of Montpellier) we consider a stochastic growth model for which extinction eventually occurs almost surely. The associated complete Fokker–Planck equation describing the law of the process is established and studied. In dimension one, e.g. for the stochastic logistic model this equation combines a PDE and an ODE (paper under revision); in dimension two, e.g. for the stochastic chemostat model this equation combines a 2D PDE and a 1D PDE [22]. We then design a finite differences numerical scheme under a probabilistic viewpoint.

6.2.3. Other modeling approaches

Participants: Anne Bisson, Jérôme Harmand, Alain Rapaport.

A collaboration with the UMR Eco & Sols has led to the development of a (static) probabilistic model for inferring nature and number of interactions in communities assembly [29]. This model has brought new insights on a data set from reconstituted soil ecosystems. Because of the curse of dimensionality, we have begun this year to extend this approach to "assembling motifs" instead of considering all the possible assemblages (paper in preparation).

In [28], ecological trade-offs between species are studied to explain species coexistence in ecological communities. In our model, plant species compete for sites where each site has a fixed stress condition. Species differ both in stress tolerance and competitive ability. We derive the deterministic discrete-time dynamical system for the species abundances. We prove the conditions under which plant species can coexist in an stable equilibrium. We compare our model with a recently proposed, continuous-time dynamical system for a tolerance-fecundity trade-off in plant communities, and we show that this model is a special case of the continuous-time version of our model.

6.3. Analysis and supervision of bioprocesses

6.3.1. Models development and identification

Participants: Fabien Campillo, Amine Charfi, Yessmine Daoud, Jérôme Harmand, Sonia Hassam, Guilherme Pimentel, Alain Rapaport.

Membrane bioreactors combine a filtration process (with a membrane) and a suspended growth rate bioreactor. This recent technology present many advantages compared to conventional ones, but is more sophisticated and requires refined control because of the fouling process. We have proposed new modeling approaches of such bioreactors, where a fouling mechanism is explicitly described:

- Membrane bioreactors allow a perfect separation of biomass and treated waters. However, membranes are subject to clogging - also called fouling - by large organic molecules and solids (biomass, suspended solids, etc...). This phenomenon represents the main problem that limits the development of membranes bioreactors. It is well documented but very studies tried to formalize it under the form of dynamical model, even more under anaerobic conditions. This is what we did in the framework of Amine Charfi's PhD [11]: he contributed to a better understanding of membrane fouling dynamics and proposed a number of strategies to avoid irreversible clogging [24], [25].
- In the scope of the PhD supervision of Guilherme Pimentel, we have proposed a simple three time scales model in view of the control of the cake formation [13], [37]. This model has been validated on real data from a pilot plant at Univ. Mons (Belgium).

As already mentioned in 6.2.1 , anaerobic bioreactors are able to produce valuable energy. However, they are subject to destabilization in case of organic overload. It is thus necessary to develop appropriate models dedicated to the synthesis of stabilizing control feedbacks. Two strategies are followed to obtain such models. Either simple models are proposed from the knowledge we have about the process, either reduced models are obtained from more complex ones. Yessmine Daoud just began her PhD thesis following the first strategy: more precisely, we study simple inhibition models of the anaerobic digestion and tries to establish simple equivalence between these models and the well known ADM1 model. Sonia Hassam, in her PhD, works within the framework of the second strategy to propose simple models obtained in reducing complex ones like the ADM1.

As already proposed last year, we have set a methodology to identify from data observed on a chemostat plant a multi-specific model that suits better than a mono-specific one, when data from molecular biology are available. In [35], we combine molecular fingerprints obtained at some discrete times (such as the ones provides by the DGGE or SSCP techniques) with on-line macroscopic measurements. In a similar spirit, within the framework of a collaboration with the LOMIC at Banyuls, we have analysed molecular data obtained via SSCP technique to monitor the structure of microbial communities. It was shown that aerial transport of bacteria from desert into the sea influenced its bacterial diversity [32], [31].

As far as stochastic approaches are concerned, the thesis of Mohsen Chebbi aims at developing stochastic models of membrane bioreactors following the approach proposed in [57]. A mathematical and simulation framework has been established, as well as the bases of vectorial simulation techniques in `Matlab`. Developments of Monte Carlo techniques for the identification of bioprocesses are investigated in the thesis of Oussama Hadj-Abdelkader, improving classical particle filtering approaches (sequential Monte Carlo) by integrating MCMC (Monte Carlo Markov Chain) procedures. A software approach has been adopted in `C++`.

6.3.2. Synthesis of control laws

Participants: Térence Bayen, Walid Bouhafs, Céline Casenave, Amel Ghouali, Jérôme Harmand, Zeyneb Khedim, Claude Lobry, Alain Rapaport, Victor Riqueleme, Matthieu Sebbah.

We investigate two kinds of bioprocesses to be controlled, arising in industrial biotechnology (digesters, waste-water purification...) or in the bioremediation of natural environments (lakes, landfill...).

6.3.2.1. Industrial biotechnology

As it is often the case in industry, we distinguish two kinds of process operating: continuous processes, for which the volumes of the bio-reactors are constant, and fed-batch processes, for which the filling rate is the control.

We tackle several optimal control problems related to the maximization of productivity of continuous bioprocesses:

- As far as anaerobic digestion is concerned, we have considered an optimal control problem for the chemostat model with substrate inhibition. The originality of this problem relies on the fact that the quantity to be maximized is not simply a state of the model (the substrate or the biomass) but the volume of biogas produced (which is a nonlinear function of the state) within a given time interval. Optimal strategies have been proposed for a class of initial conditions of the system (PhD thesis of Amel Ghouali). Other optimal control problems are studies by Walid Bouhafs to establish the optimality of controls initially proposed by Djalel Mazouni, a former PhD candidate supervised by the team. Using a tricky projection of the problem into another state space, he has shown that the problem of optimizing the degradation of two different substrates by two antagonist bacterial communities could be solved in a very general sense as the minimization of a specific functional for a very large class of nonlinear systems [19].
- The work [15] arises in the context of selection of species (widely used in agriculture and biotech-nology in order to improve productivity). For microorganisms, the selection process can be based on genetic tools. Our methodology in [15] is to drive the competition between species in a chemostat. We consider a two species chemostat model with one limiting substrate, and our aim is to optimize the selection of the species of interest. Thanks to the Pontryagin Maximum Principle, we introduce a singular feeding strategy which allows to reach the target, and we prove that the feedback control provided by this strategy is optimal. The optimal synthesis of the problem in presence of more than two species will be investigated in a future work.
- In [38], we study the problem of minimal time for a chemostat system with one limiting substrate and one species. Given a target point, the problem consists in finding an optimal feeding strategy steering any initial condition of the system to this target. This is typically of interest whenever the input substrate concentration changes yielding in a new steady state. We consider the case where the growth rate function is of Haldane type implying the existence of a singular arc that is non-necessary admissible everywhere. We provide an optimal synthesis of the problem using tools from optimal control theory.
- The work [16] studies the coupling of a culture of micro-algae limited by light and an anaerobic digester in a two-tank bioreactor (the model combines a periodic day-night light for the culture of micro-algae and a classical chemostat model for the digester). We first prove the existence and attraction of periodic solutions of this problem for a one day period. Then, we study the optimal control problem of optimizing the production of methane in the digester during a certain time frame, the control on the system being the dilution rate (the input flow of micro-algae in the digester). We also investigate the dependence of the optimal cost with respect to the volume ratio of the two tanks.

In collaboration with researchers of the unit SPO (Sciences For Oenology), we have proposed a control law of a multi-stage continuous fermentor (MSCF) designed for the study of the wine fermentation, that has been implemented on the experimental process [41]. We have also finalized the controllability analysis and minimal time feedback synthesis of models of cascade of continuous bioreactors under input constraints [18], [40] (that is also related to the control of MSCF).

As far as fed-batch processes are concerned, an extension of former results of the team about the minimal time control of fed-batch processes with impulse controls is presented in [34],

The paper [17] is devoted to the study of the minimal time problem of a fed-batch reactor, under the presence of a saturation point on the singular locus (this typically occurs whenever the growth rate function is of Haldane type and when typically the maximum input flow rate is not high enough to maintain the substrate concentration constant). This brings non-intuitive issues for the optimal synthesis (existence of switching curve and point of prior saturation).

6.3.2.2. Bioremediation of natural environments

In the scope of the associated team with Chile and the supervision of the postdoctoral stay of Matthieu Sebbah in Chile, we have addressed a new model of landfill remediation when controlling the leachate recirculation [53]. We have applied the same methodological approach than the one for the work [18], [40] mentioned in the previous section, which consists in characterizing first the sub-domains for which the target can be optimally reached with a constant extreme control (no recirculation or maximal speed of recirculation), and further the nature of optimal commutations outside these sets. This analysis provides information for the practitioners on the benefit to implement sensors and real-time controllers.

Also in the scope of the associated team with Chile (see 7.3.2.1) and the co-supervision of the PhD of Victor Riquelme, we have carried on the study of optimal syntheses for the minimal time treatment of natural water reservoirs (such as lakes) [52]. We have proved that the minimal time strategy consists in a most-rapid approach to homogeneous concentrations, even though the optimal control problem is non convex. Moreover, we have shown that a large diffusion increases the treatment time when the resource is everywhere highly polluted , while it can at the opposite be beneficial when only part of the resource is polluted. This feature should serve the practitioners in the choice of pumps positioning in a originally clean water resource that is suddenly affected by a local pollution. This work is in connection with the INRA/Inria patent [47] that has been deposited jointly with LEMON Team.

6.4. Other application domains

Participants: Fabien Campillo, Céline Casenave.

Semi-Markov land use dynamic

With IRD (GRED Montpellier) and the Univ. of Fianarantsoa (Madagascar) we pursued our study on land use dynamics models corresponding to parcels located on the edge of the forest corridor, Madagascar. We use semi-Markov chain to infer the land-use dynamics. In addition to the empirical and maximum likelihood methods, we estimate the semi-Markov kernel by a Bayesian approach [21].

Ice cream crystallization

We study the problem of the control of an ice cream crystallization process, part of the European CAFE project, in collaboration with CESAME (Univ. Catholique de Louvain-la-neuve), Irstea Antony and AgroParisTech. The goal is to control the viscosity of the ice cream at the outlet of the continuous crystallizer. On the basis of a population balance equation describing the evolution of the crystal size distribution of the ice cream, and an energy balance equation, we have proposed an input-output reduced order model of the process, that has been identified and validated on experimental data [23]. A nonlinear control strategy based on an adaptive linearizing control law coupled with a Smith predictor to account for the measurement delay has been proposed, and validated on the experimental pilot plant (paper in preparation).

MOISE Project-Team

6. New Results

6.1. Mathematical Modelling of the Ocean Dynamics

6.1.1. Numerical Schemes for Ocean Modelling

Participants: Eric Blayo, Laurent Debreu, Jérémie Demange, Florian Lemarié.

In his PhD, Jérémie Demange has worked on advection-diffusion schemes for ocean models (Supervisors : L. Debreu, P. Marchesiello (IRD)). His work focuses on the link between tracers (temperature and salinity) and momentum advection and diffusion in the non hyperbolic system of equations typically used in ocean models (the so called primitive equations with hydrostatic and Boussinesq assumptions). We also investigated the use of a depth dependent barotropic mode in free surface ocean models. When most ocean models assume that this mode is vertically constant, we have shown that the use of the true barotropic mode, derived from a normal mode decomposition, allows more stability and accuracy in the representation of external gravity waves ([55]). A special focus has also been put on the numerical representation of internal gravity waves (IGW). The normal mode decomposition also allows the computation of IGW characteristic variables and speeds and thus enables the derivation of monotonic advection schemes ([54]).

In 2014, we worked on the stability constraints for oceanic numerical models ([56]). The idea is to carry a deep analysis of these constraints in order to propose new time stepping algorithms for ocean models. Except for vertical diffusion (and possibly the external mode and bottom drag), oceanic models usually rely on explicit time-stepping algorithms subject to Courant-Friedrichs-Lewy (CFL) stability criteria. Implicit methods could be unconditionally stable, but an algebraic system must be solved at each time step and other considerations such as accuracy and efficiency are less straightforward to achieve. Depending on the target application, the process limiting the maximum allowed time-step is generally different. In this paper, we introduce offline diagnostics to predict stability limits associated with internal gravity waves, advection, diffusion, and rotation. This suite of diagnostics is applied to a set of global, regional and coastal numerical simulations with several horizontal/vertical resolutions and different numerical models. We show that, for resolutions finer that 1/2°, models with an Eulerian vertical coordinate are generally constrained by vertical advection in a few hot spots and that numerics must be extremely robust to changes in Courant number. Based on those results, we review the stability and accuracy of existing numerical kernels in vogue in primitive equations oceanic models with a focus on advective processes and the dynamics of internal waves. We emphasize the additional value of studying the numerical kernel of oceanic models in the light of coupled space-time approaches instead of studying the time schemes independently from spatial discretizations. From this study, we suggest some guidelines for the development of temporal schemes in future generation multi-purpose oceanic models.

6.1.2. Coupling Methods for Oceanic and Atmospheric Models

Participants: Eric Blayo, Mehdi-Pierre Daou, Laurent Debreu, Florian Lemarié, Antoine Rousseau.

6.1.2.1. Coupling dimensionally heterogeneous models

The coupling of different types of models is gaining more and more attention recently. This is due, in particular, to the needs of more global models encompassing different disciplines (e.g. multi-physics) and different approaches (e.g. multi-scale, nesting). Also, the possibility to assemble different modeling units inside a friendly modelling software platform is an attractive solution compared to developing more and more complex global models. More specifically one may want to couple 1D to 2D or 3D models, such as Shallow Water and Navier Stokes models: this was the framework of our partnership with EDF, now extended with ARTELIA Group.

Following the work done by Manel Tayachi in her PhD, Medhi Pierre Daou has started implementing and analyzing a coupling between 1D shallow water equations and 3D Navier Stokes equations. In the context of our partnership with ARTELIA, he uses industrial codes (Mascaret, Telemac and OpenFoam). A first implementation has been realized in an academic testcase, and a second one is presently under implementation in a much more realistic context, in the framework of the European project CRISMA.

6.1.2.2. Ocean-atmosphere coupling

Coupling methods routinely used in regional and global climate models do not provide the exact solution to the ocean-atmosphere problem, but an approached one [57]. For the last few years we have been actively working on the analysis of Schwarz waveform relaxation to apply this type of iterative coupling method to air-sea coupling [95], [96], [94]. In the context of the simulation of tropical cyclone, sensitivity tests to the coupling method have been carried out in an ensemblist approach. We showed that with a mathematically consistent coupling, compared to coupling methods en vogue in existing coupled models, the spread of the ensemble is reduced, thus indicating a much reduced uncertainty in the physical solution. In 2014, this work has been the subject of several invited conferences [23], [24], [25], [26] and collaborations with geophysicists [41], [47], [48].

Past year has also been dedicated to the establishment of strong collaborations between the applied mathematics and the climate community to assess the impact of our work on IPCC-like climate models and to go further in the theoretical work by including the formulation of physical parameterizations. As a results, a PhD-thesis (C. Pelletier) funded by Inria has started in fall 2014 in collaboration with the LSCE (Laboratoire des Sciences du Climat et de l'Environnement). Moreover a PPR (*Projet à partenariat renforcé*) called SIMBAD (SIMplified Boundary Atmospheric layer moDel for ocean modeling purposes) is funded by Mercator-Ocean for the next three years. The aim of this project in collaboration with Meteo-France, Ifremer, LMD, and LOCEAN is to derive a metamodel to force high-resolution oceanic operational models for which the use of a full atmospheric model is not possible due to a prohibitive computational cost.

6.2. Development of New Methods for Data Assimilation

6.2.1. Multigrid Methods for Variational Data Assimilation.

Participants: Laurent Debreu, François-Xavier Le Dimet, Arthur Vidard.

In order to lower the computational cost of the variational data assimilation process, we investigate the use of multigrid methods to solve the associated optimal control system. On a linear advection equation, we study the impact of the regularization term on the optimal control and the impact of discretization errors on the efficiency of the coarse grid correction step. We show that even if the optimal control problem leads to the solution of an elliptic system, numerical errors introduced by the discretization can alter the success of the multigrid methods. The view of the multigrid iteration as a preconditioner for a Krylov optimization method leads to a more robust algorithm. A scale dependent weighting of the multigrid preconditioner and the usual background error covariance matrix based preconditioner is proposed and brings significant improvements. This work is presented in a paper submitted to QJRMS ([84]). A book chapter on multiresolution methods for data assimilation has also been published ([51]).

6.2.2. Variational Data Assimilation with Control of Model Error

Participant: Arthur Vidard.

One of the main limitations of the current operational variational data assimilation techniques is that they assume the model to be perfect, mainly because of computing cost issues. Numerous researches have been carried out to reduce the cost of controlling model errors by controlling the correction term only in certain privileged directions or by controlling only the systematic and time correlated part of the error.

Both the above methods consider the model errors as a forcing term in the model equations. Trémolet (2006) describes another approach where the full state vector (4D field: 3D spatial + time) is controlled. Because of computing cost one cannot obviously control the model state at each time step. Therefore, the assimilation window is split into sub-windows, and only the initial conditions of each sub-window are controlled, the junctions between each sub-window being penalized. One interesting property is that, in this case, the computation of the gradients, for the different sub-windows, are independent and therefore can be done in parallel.

This method is now implemented in a realistic oceanic framework using OPAVAR/ NEMOVAR. The plan is to extend this study focusing on the parallel aspects of such approach.

6.2.3. Assimilation of Images

Participants: François-Xavier Le Dimet, Maëlle Nodet, Arthur Vidard, Nelson Feyeux, Vincent Chabot, Nicolas Papadakis.

6.2.3.1. Direct assimilation of image sequences

At the present time the observation of Earth from space is done by more than thirty satellites. These platforms provide two kinds of observational information:

- Eulerian information as radiance measurements: the radiative properties of the earth and its fluid envelops. These data can be plugged into numerical models by solving some inverse problems.
- Lagrangian information: the movement of fronts and vortices give information on the dynamics of the fluid. Presently this information is scarcely used in meteorology by following small cumulus clouds and using them as Lagrangian tracers, but the selection of these clouds must be done by hand and the altitude of the selected clouds must be known. This is done by using the temperature of the top of the cloud.

MOISE was the leader of the ANR ADDISA project dedicated to the assimilation of images, and is a member of its follow-up GeoFluids (along with EPI FLUMINANCE and CLIME, and LMD, IFREMER and Météo-France) that just ended in 2013.

During the ADDISA project we developed Direct Image Sequences Assimilation (DISA) and proposed a new scheme for the regularization of optical flow problems [102]. Thanks to the nonlinear brightness assumption, we proposed an algorithm to estimate the motion between two images, based on the minimization of a nonlinear cost function. We proved its efficiency and robustness on simulated and experimental geophysical flows [76]. As part of the ANR project GeoFluids, we are investigating new ways to define distance between a couple of images. One idea is to compare the gradient of the images rather than the actual value of the pixels. This leads to promising results. Another idea, currently under investigation, consists in comparing main structures within each image. This can be done using, for example, a wavelet representation of images. Both approaches have been compared, in particular their relative merits in dealing with observation errors, in a paper accepted late 2014 [8] and presented in several national [37], [38], [27] and international conferences [30], [28], [29].

Vincent Chabot also defended his PhD in July 2014 [2].

In recent developments [17] we have also used "Level Sets" methods to describe the evolution of the images. The advantage of this approach is that it permits, thanks to the level sets function, to consider the images as a state variable of the problem. We have derived an Optimality System including the level sets of the images.

6.2.3.2. Optimal transport for image assimilation

Within the optimal transport project TOMMI funded by the ANR white program (started mid 2011), a new optimization scheme based on proximal splitting method has been proposed to solve the dynamic optimal transport problem. We investigate the use of optimal transport based distances for data assimilation. N. Feyeux started his PhD on this subject last year. The study is still under investigation, but preliminary encouraging results have already been presented twice, in France [68] and Austria [69].

6.2.4. A Nudging-Based Data Assimilation Method: the Back and Forth Nudging

Participants: Maëlle Nodet, Jacques Blum, Didier Auroux.

The Back and Forth Nudging (BFN) algorithm has been recently introduced for simplicity reasons, as it does not require any linearization, nor adjoint equation, or minimization process in comparison with variational schemes. Nevertheless it provides a new estimation of the initial condition at each iteration.

Previous theoretical results showed that BFN was often ill-posed for viscous partial differential equations. To overcome this problem, we proposed a new version of the algorithm, which we called the Diffusive BFN, and which showed very promising results on one-dimensional viscous equations. Experiments on more sophisticated geophysical models, such as Shallow-Water equations and NEMO ocean model are still in progress, in collaboration with University of Nice, and have been presented at the ICIPE conference [31].

6.2.5. Variational Data Assimilation and Control of Boundary Conditions

Participant: Eugène Kazantsev.

A variational data assimilation technique is applied to the identification of the optimal boundary conditions for a simplified configuration of the NEMO model. A rectangular box model placed in mid-latitudes, and subject to the classical single or double gyre wind forcing, is studied. The model grid can be rotated on a desired angle around the center of the rectangle in order to simulate the boundary approximated by a staircase-like coastlines. The solution of the model on the grid aligned with the box borders was used as a reference solution and as artificial observational data. It is shown in [15] that optimal boundary has a rather complicated geometry which is neither a staircase, nor a straight line. The boundary conditions found in the data assimilation procedure bring the solution toward the reference solution allowing to correct the influence of the rotated grid (see fig. 1).

Adjoint models, necessary to variational data assimilation, have been produced by the TAPENADE software, developed by the SCIPORT team. This software is shown to be able to produce the adjoint code, that can be used in data assimilation after a memory usage optimization.

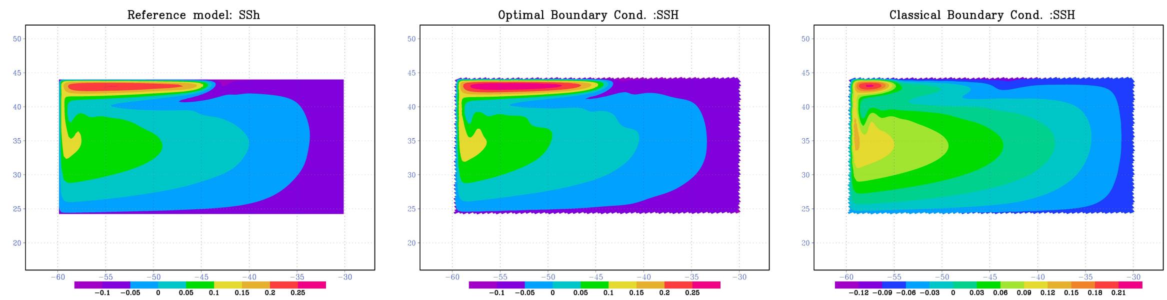


Figure 1. Sea surface elevation: reference solution on the aligned grid (left), solutions on the 30° rotated grid with optimal (center) and classical (right) boundary conditions.

6.3. Data Assimilation for Geophysical Models

6.3.1. Development of a Variational Data Assimilation System for OPA9/NEMO

Participants: Arthur Vidard, Pierre-Antoine Bouttier, Eric Blayo.

We are heavily involved in the development of NEMOVAR (Variational assimilation for NEMO). For several years now, we built a working group (coordinated by A. Vidard) in order to bring together various NEMOVAR user-groups with diverse scientific interests (ranging from singular vector and sensitivity studies to specific issues in variational assimilation). It has led to the creation of the VODA (Variational Ocean Data Assimilation for multi scales applications) ANR project (ended in 2012). A new project, part of a larger EU-FP7 project (ERA-CLIM2) has just started in january 2014.
The project aims at delivering a common NEMOVAR platform based on NEMO platform for 3D and 4D variational assimilation. Following 2009-11 VODA activities, a fully parallel version of NEMOTAM (Tangent and Adjoint Model for NEMO) is now available for the community in the standard NEMO version. This version is based on the released 3.4.1 version of NEMO.

We are also investigating variational data assimilation methods applied to high resolution ocean numerical models (see figure 2). This part of the project is now well advanced and encouraging preliminary results are available on an idealised numerical configuration of an oceanic basin. Several novative diagnostics have been also developed in this framework as part of P.A. Bouttier's PhD that was defended early 2014 [1].

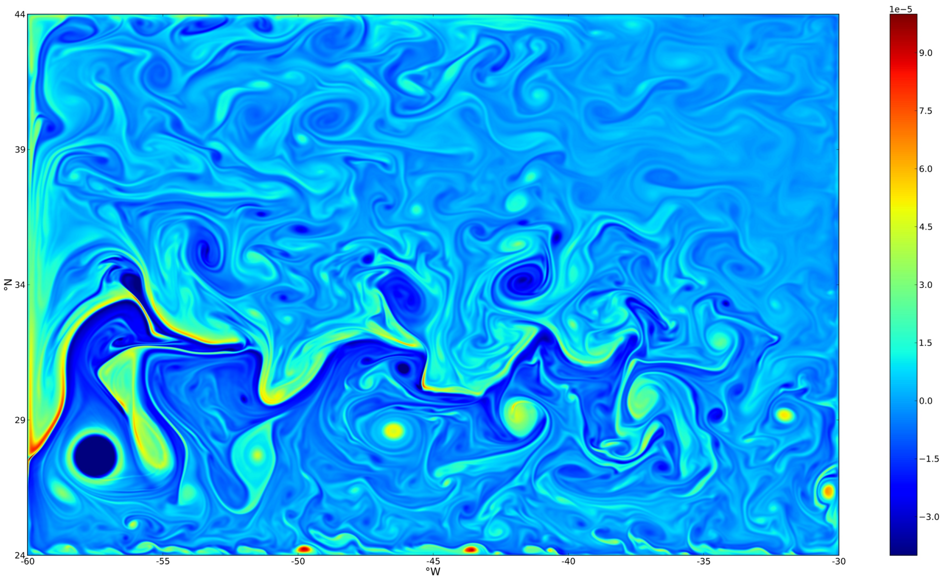


Figure 2. Snapshot of the relative vorticity field (1/s) for an academic oceanic basin model at 1/100° horizontal resolution.

Lastly, multi resolution algorithms have been developed to solve the variational problem. An EU-ITN (International Training Network) project is going to be submitted early 2015 to continue working in this particular aspect.

6.3.2. Ensemble Kalman Filtering for Large Scale Ice-Sheet Models

Participants: Bertrand Bonan, Maëlle Nodet, Catherine Ritz.

In collaboration with C. Ritz (CNRS, Laboratoire de Glaciologie et Geophysique de l'Environnement (LGGE), Grenoble), we aim to develop inverse methods for ice cap models.

In the framework of global warming, the evolution of sea level is a major but ill-known phenomenon. It is difficult to validate the models which are used to predict the sea level elevation, because observations are heterogeneous and sparse.

Data acquisition in polar glaciology is difficult and expensive. Satellite data have a good spatial coverage, but they allow only indirect observation of the interesting data. Moreover, ice dynamics processes are highly non linear and involve many feedback loops, so that classical linear data assimilation gives poor results.

B. Bonan defended his PhD in November 2013 on this subject. We implemented the Ensemble Transform Kalman Filter (ETKF) algorithm for a flowline Shallow-Ice model, called Winnie, developed by C. Ritz at LGGE. On twin experiments we got interesting results, very promising for the future, as we want to implement this method into a full 3D model. A journal paper has published on this subject [5], and the results have been presented in the conference [46].

6.4. Quantifying Uncertainty

6.4.1. Sensitivity Analysis for Forecasting Ocean Models

Participants: Eric Blayo, Laurent Gilquin, Céline Helbert, François-Xavier Le Dimet, Simon Nanty, Maëlle Nodet, Clémentine Prieur, Laurence Viry, Federico Zertuche.

6.4.1.1. Scientific context

Forecasting geophysical systems require complex models, which sometimes need to be coupled, and which make use of data assimilation. The objective of this project is, for a given output of such a system, to identify the most influential parameters, and to evaluate the effect of uncertainty in input parameters on model output. Existing stochastic tools are not well suited for high dimension problems (in particular time-dependent problems), while deterministic tools are fully applicable but only provide limited information. So the challenge is to gather expertise on one hand on numerical approximation and control of Partial Differential Equations, and on the other hand on stochastic methods for sensitivity analysis, in order to develop and design innovative stochastic solutions to study high dimension models and to propose new hybrid approaches combining the stochastic and deterministic methods.

6.4.1.2. Estimating sensitivity indices

A first task is to develop tools for estimated sensitivity indices. In variance-based sensitivity analysis, a classical tool is the method of Sobol' [101] which allows to compute Sobol' indices using Monte Carlo integration. One of the main drawbacks of this approach is that the estimation of Sobol' indices requires the use of several samples. For example, in a d-dimensional space, the estimation of all the first-order Sobol' indices requires $d+1$ samples. Some interesting combinatorial results have been introduced to weaken this defect, in particular by Saltelli [99] and more recently by Owen [97] but the quantities they estimate still require $O(d)$ samples.

In a recent work [21] we introduce a new approach to estimate all first-order Sobol' indices by using only two samples based on replicated latin hypercubes and all second-order Sobol' indices by using only two samples based on replicated randomized orthogonal arrays. We establish theoretical properties of such a method for the first-order Sobol' indices and discuss the generalization to higher-order indices. As an illustration, we propose to apply this new approach to a marine ecosystem model of the Ligurian sea (northwestern Mediterranean) in order to study the relative importance of its several parameters. The calibration process of this kind of chemical simulators is well-known to be quite intricate, and a rigorous and robust — i.e. valid without strong regularity assumptions — sensitivity analysis, as the method of Sobol' provides, could be of great help. The computations are performed by using CIGRI, the middleware used on the grid of the Grenoble University High Performance Computing (HPC) center. We are also applying these estimates to calibrate integrated land use transport models. As for these models, some groups of inputs are correlated, Laurent Gilquin extended the approach based on replicated designs for the estimation of grouped Sobol' indices [70].

We can now wonder what are the asymptotic properties of these new estimators, or also of more classical ones. In [10], the authors deal with asymptotic properties of the estimators. In [89], the authors establish also a multivariate central limit theorem and non asymptotic properties.

6.4.1.3. Intrusive sensitivity analysis, reduced models

Another point developed in the team for sensitivity analysis is model reduction. To be more precise regarding model reduction, the aim is to reduce the number of unknown variables (to be computed by the model), using a well chosen basis. Instead of discretizing the model over a huge grid (with millions of points), the state vector of the model is projected on the subspace spanned by this basis (of a far lesser dimension). The choice of the basis is of course crucial and implies the success or failure of the reduced model. Various model reduction methods offer various choices of basis functions. A well-known method is called “proper orthogonal decomposition" or “principal component analysis". More recent and sophisticated methods also exist and may be studied, depending on the needs raised by the theoretical study. Model reduction is a natural way to overcome difficulties due to huge computational times due to discretizations on fine grids. In [92], the authors present a reduced basis offline/online procedure for viscous Burgers initial boundary value problem, enabling efficient approximate computation of the solutions of this equation for parametrized viscosity and initial and boundary value data. This procedure comes with a fast-evaluated rigorous error bound certifying the approximation procedure. The numerical experiments in the paper show significant computational savings, as well as efficiency of the error bound.

When a metamodel is used (for example reduced basis metamodel, but also kriging, regression, ...) for estimating sensitivity indices by Monte Carlo type estimation, a twofold error appears: a sampling error and a metamodel error. Deriving confidence intervals taking into account these two sources of uncertainties is of great interest. We obtained results particularly well fitted for reduced basis metamodels [14]. In [91], the authors provide asymptotic confidence intervals in the double limit where the sample size goes to infinity and the metamodel converges to the true model. These results were also adapted to problems related to more general models such as Shallow-Water equations, in the context of the control of an open channel [72].

Let us come back to the output of interest. Is it possible to get better error certification when the output is specified. A work in this sense has been submitted, dealing with goal oriented uncertainties assessment [71].

6.4.1.4. Sensitivity analysis with dependent inputs

An important challenge for stochastic sensitivity analysis is to develop methodologies which work for dependent inputs. For the moment, there does not exist conclusive results in that direction. Our aim is to define an analogue of Hoeffding decomposition [90] in the case where input parameters are correlated. Clémentine Prieur supervised Gaëlle Chastaing's PhD thesis on the topic (defended in September 2013) [78]. We obtained first results [79], deriving a general functional ANOVA for dependent inputs, allowing defining new variance based sensitivity indices for correlated inputs. We then adapted various algorithms for the estimation of these new indices. These algorithms make the assumption that among the potential interactions, only few are significant. Two papers have been recently accepted [66] and [80]. We also considered (see the paragraph 6.4.1) the estimation of groups Sobol' indices, with a procedure based on replicated designs. These indices provide information at the level of groups, and not at a finer level, but their interpretation is still rigorous.

Céline Helbert and Clémentine Prieur supervise the PhD thesis of Simon Nanty (funded by CEA Cadarache). The subject of the thesis is the analysis of uncertainties for numerical codes with temporal and spatio-temporal input variables, with application to safety and impact calculation studies. This study implies functional dependent inputs. A first step is the modeling of these inputs, and a paper has been submitted [74].

6.4.1.5. Multy-fidelity modeling for risk analysis

Federico Zertuche's PhD concerns the modeling and prediction of a digital output from a computer code when multiple levels of fidelity of the code are available. A low-fidelity output can be obtained, for example on a coarse mesh. It is cheaper, but also much less accurate than a high-fidelity output obtained on a fine mesh. In this context, we propose new approaches to relieve some restrictive assumptions of existing methods ([93], [98]): a new estimation method of the classical cokriging model when designs are not nested and a nonparametric modeling of the relationship between low-fidelity and high-fidelity levels. The PhD takes place in the REDICE consortium and in close link with industry. The first part of the thesis was also dedicated to the development of a case study in fluid mechanics with CEA in the context of the study of a nuclear reactor.

The second part of the thesis was dedicated to the development of a new sequential approach based on a course to fine wavelets algorithm. Federico Zertuche presented his work at the annual meeting of the GDR Mascot Num in 2014 [36].

6.4.1.6. Data assimilation and second order sensitivity analysis

A main advantage of Variational Methods in Data Assimilation is to exhibit a so-called Optimality System (OS) that contains all the available information : model, data, statistics. Therefore a sensitivity analysis (i.e. the evaluation of the gradient) with respect to the inputs of the model has to be carried out on the OS. With iMECH and INM we have applied sensitivity analysis in the framework of a pollution problem in a lake. The application of second order analysis for sensitivity permits to evaluate the sensitivity with respect to observations and furthermore to determine the optimal location of new sensors at points with the highest sensitivity [16], [52].

This methodology has been applied to

- **Oil Spill**. These last years have known several disasters produced by wrecking of ships and drifting platforms with severe consequences on the physical and biological environments. In order to minimize the impact of these oil spills its necessary to predict the evolution of oil spot. Some basic models are available and some satellites provide images on the evolution of oil spots. Clearly this topic is a combination of the two previous one: data assimilation for pollution and assimilation of images. A theoretical framework has been developed with Dr. Tran Thu Ha (iMech).
- **Data Assimilation in Supercavitation (with iMech)**. Some self propelled submarine devices can reach a high speed thanks to phenomenon of supercavitation: an air bubble is created on the nose of the device and reduces drag forces. Some models of supercavitation already exist but are working on two applications of variational methods to supercavitation:
  - Parameter identification : the models have some parameters that can not be directly measured. From observations we retrieve the unknown parameters using a classical formalism of inverse problems.
  - Shape Optimization. The question is to determine an optimum design of the shape of the engine in order to reach a maximum speed.

6.5. Tracking of Mesoscale Convective Systems

Participant: Clémentine Prieur.

We are interested in the tracking of mesoscale convective systems.A particular region of interest is West Africa. Data and hydrological expertise is provided by T. Vischel and T. Lebel (LTHE, Grenoble).

A first approach involves adapting the multiple hypothesis tracking (MHT) model originally designed by the NCAR (National Centre for Atmospheric Research) for tracking storms [103] to the data for West Africa. With A. Makris (working on a post-doctoral position), we proposed a Bayesian approach [18], which consists in considering that the state at time t is composed on one hand by the events (birth, death, splitting, merging) and on the other hand by the targets' attributes (positions, velocities, sizes, ...). The model decomposes the state into two sub-states: the events and the targets positions/attributes. The events are updated first and are conditioned to the previous targets sub-state. Then given the new events the target substate is updated. A simulation study allowed to verify that this approach improves the frequentist approach by Storlie et al. (2009). It has been tested on simulations [18] and investigated in the specific context of real data on West Africa (submitted paper). Using PHD (probability hypothesis density) filters adapted to our problem, generalising recent developments in particle filtering for spatio-temporal branching processes (e.g. [77]) could be an interesting alternative to explore. The idea of a dynamic, stochastic tracking model should then provide the base for generating rainfall scenarios over a relatively vast area of West Africa in order to identify the main sources of variability in the monsoon phenomenon.

6.6. Multivariate Risk Indicators

Participants: Clémentine Prieur, Patricia Tencaliec.

Studying risks in a spatio-temporal context is a very broad field of research and one that lies at the heart of current concerns at a number of levels (hydrological risk, nuclear risk, financial risk etc.). Stochastic tools for risk analysis must be able to provide a means of determining both the intensity and probability of occurrence of damaging events such as e.g. extreme floods, earthquakes or avalanches. It is important to be able to develop effective methodologies to prevent natural hazards, including e.g. the construction of barrages.

Different risk measures have been proposed in the one-dimensional framework . The most classical ones are the return level (equivalent to the Value at Risk in finance), or the mean excess function (equivalent to the Conditional Tail Expectation CTE). However, most of the time there are multiple risk factors, whose dependence structure has to be taken into account when designing suitable risk estimators. Relatively recent regulation (such as Basel II for banks or Solvency II for insurance) has been a strong driver for the development of realistic spatio-temporal dependence models, as well as for the development of multivariate risk measurements that effectively account for these dependencies.

We refer to [81] for a review of recent extensions of the notion of return level to the multivariate framework. In the context of environmental risk, [100] proposed a generalization of the concept of return period in dimension greater than or equal to two. Michele et al. proposed in a recent study [82] to take into account the duration and not only the intensity of an event for designing what they call the dynamic return period. However, few studies address the issues of statistical inference in the multivariate context. In [86], [88], we proposed non parametric estimators of a multivariate extension of the CTE. As might be expected, the properties of these estimators deteriorate when considering extreme risk levels. In collaboration with Elena Di Bernardino (CNAM, Paris), Clémentine Prieur is working on the extrapolation of the above results to extreme risk levels.

Elena Di Bernardino, Véronique Maume-Deschamps (Univ. Lyon 1) and Clémentine Prieur also derived an estimator for bivariate tail [87]. The study of tail behavior is of great importance to assess risk.

With Anne-Catherine Favre (LTHE, Grenoble), Clémentine Prieur supervises the PhD thesis of Patricia Tencaliec. We are working on risk assessment, concerning flood data for the Durance drainage basin (France). The PhD thesis started in October 2013. A first paper on data reconstruction has been submitted. It was a necessary step as the initial series contained many missing data.

6.7. Non-Parametric Estimation for Kinetic Diffusions

Participant: Clémentine Prieur.

This research is the subject of a collaboration with Venezuela (Professor Jose R. Leon, Caracas Central University) and is partly funded by an ECOS Nord project.

We are focusing our attention on models derived from the linear Fokker-Planck equation. From a probabilistic viewpoint, these models have received particular attention in recent years, since they are a basic example for hypercoercivity. In fact, even though completely degenerated, these models are hypoelliptic and still verify some properties of coercivity, in a broad sense of the word. Such models often appear in the fields of mechanics, finance and even biology. For such models we believe it appropriate to build statistical non-parametric estimation tools. Initial results have been obtained for the estimation of invariant density, in conditions guaranteeing its existence and unicity [6] and when only partial observational data are available. A paper on the non parametric estimation of the drift has been accepted recently [7] (see Samson et al., 2012, for results for parametric models). As far as the estimation of the diffusion term is concerned, a paper has been submitted [7], in collaboration with J.R. León (Caracas, Venezuela) and P. Cattiaux (Toulouse). Recursive estimators have been also proposed by the same authors in [64] recently submitted.

Note that Professor Jose R. León (Caracas, Venezuela) is now funded by an international Inria Chair and will spend one year in our team, allowing to collaborate further on parameter estimation.

6.8. Land Use and Transport Models Calibration

Participants: Thomas Capelle, Laurent Gilquin, Clémentine Prieur, Arthur Vidard, Peter Sturm, Elise Arnaud.

Given the complexity of modern urban areas, designing sustainable policies calls for more than sheer expert knowledge. This is especially true of transport or land use policies, because of the strong interplay between the land use and the transportation systems. Land use and transport integrated (LUTI) modelling offers invaluable analysis tools for planners working on transportation and urban projects. Yet, very few local authorities in charge of planning make use of these strategic models. The explanation lies first in the difficulty to calibrate these models, second in the lack of confidence in their results, which itself stems from the absence of any well-defined validation procedure. Our expertise in such matters will probably be valuable for improving the reliability of these models. To that purpose we participated to the building up of the ANR project CITiES led by the STEEP EPI. This project started early 2013 and two PhD about sensitivity analysis and calibration were launched late 2013. This work led to conference papers [32], [33] and a submitted journal paper [70]

MOKAPLAN Team

6. New Results

6.1. Highlights of the Year

All of the new results below are important break through and most of them non-incremental research.

Mokaplan has extended its collaborations to several researchers at Ceremade and is under review to become a project team.

6.2. Iterative Bregman Projections for Regularized Transportation Problems

Benamou, Jean-David and Carlier, Guillaume and Cuturi, Marco and Nenna, Luca and Peyré, Gabriel

[19]

We provide a general numerical framework to approximate solutions to linear programs related to optimal transport. The general idea is to introduce an entropic regularization of the initial linear program. This regularized problem corresponds to a Kullback-Leibler Bregman di-vergence projection of a vector (representing some initial joint distribution) on the polytope of constraints. We show that for many problems related to optimal transport, the set of linear constraints can be split in an intersection of a few simple constraints, for which the projections can be computed in closed form. This allows us to make use of iterative Bregman projections (when there are only equality constraints) or more generally Bregman-Dykstra iterations (when inequality constraints are involved). We illustrate the usefulness of this approach to several variational problems related to optimal transport: barycenters for the optimal trans-port metric, tomographic reconstruction, multi-marginal optimal trans-port and in particular its application to Brenier's relaxed solutions of in-compressible Euler equations, partial unbalanced optimal transport and optimal transport with capacity constraints.
The extension of the method to the Principal Agent problem, Density Functional theory and Transport under martingal constraint is under way.

6.3. A viscosity framework for computing Pogorelov solutions of the Monge-Ampere equation

Benamou, Jean-David and Froese, Brittany D.

[21]

We consider the Monge-Kantorovich optimal transportation problem between two measures, one of which is a weighted sum of Diracs. This problem is traditionally solved using expensive geometric methods. It can also be reformulated as an elliptic partial differential equation known as the Monge-Ampere equation. However, existing numerical methods for this non-linear PDE require the measures to have finite density. We introduce a new formulation that couples the viscosity and Aleksandrov solution definitions and show that it is equivalent to the original problem. Moreover, we describe a local reformulation of the subgradient measure at the Diracs, which makes use of one-sided directional derivatives. This leads to a consistent, monotone discretisation of the equation. Computational results demonstrate the correctness of this scheme when methods designed for conventional viscosity solutions fail.
The method offers a new insight into the duality between Aleksandrov and Brenier solutions of the Monge Ampère equations. We still work on the viscosity existence/uniqueness convergence of sheme theory.

6.4. Discretization of functionals involving the Monge-Ampère operator

Benamou, Jean-David and Carlier, Guillaume and Mérigot, Quentin and Oudet, Edouard

[26]

Gradient flows in the Wasserstein space have become a powerful tool in the analysis of diffusion equations, following the seminal work of Jordan, Kinderlehrer and Otto (JKO). The numerical applications of this formulation have been limited by the difficulty to compute the Wasserstein distance in dimension larger than 2. One step of the JKO scheme is equivalent to a variational problem on the space of convex functions, which involves the Monge-Ampère operator. Convexity constraints are notably difficult to handle numerically, but in our setting the internal energy plays the role of a barrier for these constraints. This enables us to introduce a consistent discretization, which inherits convexity properties of the continuous variational problem. We show the effectiveness of our approach on nonlinear diffusion and crowd-motion models.

6.5. Augmented Lagrangian methods for transport optimization, Mean-Field Games and degenerate PDEs

Benamou, Jean-David and Carlier, Guillaume

[18]

Many problems from mass transport can be reformulated as variational problems under a prescribed divergence constraint (static problems) or subject to a time dependent continuity equation which again can also be formulated as a divergence constraint but in time and space. The variational class of Mean-Field Games introduced by Lasry and Lions may also be interpreted as a generalisation of the time-dependent optimal transport problem. Following Benamou and Brenier, we show that augmented Lagrangian methods are well-suited to treat convex but nonsmooth problems. It includes in particular Monge historic optimal transport problem. A Finite Element discretization and implementation of the method is used to provide numerical simulations and a convergence study.
We have good hopes to use this method to many non-linear diffusion equations through the use of JKO gradient schemes.

6.6. Discretization of functionals involving the Monge-Ampère operator

Benamou, Jean-David and Collino, Francis and Mirebeau, Jean-Marie

[20]

We introduce a novel discretization of the Monge-Ampere operator, simultaneously consistent and degenerate elliptic, hence accurate and robust in applications. These properties are achieved by exploiting the arithmetic structure of the discrete domain, assumed to be a two dimensional cartesian grid. The construction of our scheme is simple, but its analysis relies on original tools seldom encountered in numerical analysis, such as the geometry of two dimensional lattices, and an arithmetic structure called the Stern-Brocot tree. Numerical experiments illustrate the method's efficiency.

6.7. A Γ-Convergence Result for the Upper Bound Limit Analysis of Plates

Bleyer, Jérémy and Carlier, Guillaume and Duval, Vincent and Mirebeau, Jean-Marie and Peyré, Gabriel

[23]

Upper bound limit analysis allows one to evaluate directly the ultimate load of structures without performing a cumbersome incremental analysis. In order to numerically apply this method to thin plates in bending, several authors have proposed to use various finite elements discretizations. We provide in this paper a mathematical analysis which ensures the convergence of the finite element method, even with finite elements with discontinuous derivatives such as the quadratic 6 node Lagrange triangles and the cubic Hermite triangles. More precisely, we prove the Gamma-convergence of the discretized problems towards the continuous limit analysis problem. Numerical results illustrate the relevance of this analysis for the yield design of both homogeneous and non-homogeneous materials.

6.8. Cournot-Nash equilibria

Carlier, Guillaume and Blanchet, Adrien

[24]

The notion of Nash equilibria plays a key role in the analysis of strategic interactions in the framework of N player games. Analysis of Nash equilibria is however a complex issue when the number of players is large. It is therefore natural to investigate the continuous limit as N tends to infinity and to investigate whether it corresponds to the notion of Cournot-Nash equilibria. In [9], this kind of convergence result is studied in a Wasserstein framework. In [BC1], we go one step further by giving a class of games with a continnum of players for which equilibria may be found as minimizers as a functional on measures which is very similar to the one-step JKO case, uniqueness results are the obtained from displacement convexity arguments. Finally, in [9] some situations which are non variational are considered and existence is obtained by methods combining fixed point arguments and optimal transport.

6.9. Principal Agent

Carlier, Guillaume, Benamou, Jean-David and Dupuis Xavier

The numerical resolution of principal Agent for a bilinear utility has been attacked and solved successfully in a series of recent papers see [70] and references therein.

A Bregman approach inspired by [6] has been developed for more general functions the paper is currently being written. It would be extremely useful as a complement to the theoretical analysis. A new semi-Discrete Geometric approach is also investigated where the method reduces to non-convex polynomial optimization.

6.10. Exact Support Recovery for Sparse Spikes Deconvolution

Duval, Vincent and Peyré, Gabriel

[17]

We study sparse spikes deconvolution over the space of measures. We focus our attention to the recovery properties of the support of the measure, i.e. the location of the Dirac masses. For non-degenerate sums of Diracs, we show that, when the signal-to-noise ratio is large enough, total variation regularization (which is the natural extension of the L1 norm of vectors to the setting of measures) recovers the exact same number of Diracs. We also show that both the locations and the heights of these Diracs converge toward those of the input measure when the noise drops to zero. The exact speed of convergence is governed by a specific dual certificate, which can be computed by solving a linear system. We draw connections between the support of the recovered measure on a continuous domain and on a discretized grid. We show that when the signal-to-noise level is large enough, the solution of the discretized problem is supported on pairs of Diracs which are neighbors of the Diracs of the input measure. This gives a precise description of the convergence of the solution of the discretized problem toward the solution of the continuous grid-free problem, as the grid size tends to zero.

MORPHEME Project-Team

5. New Results

5.1. Highlights of the Year

- Laure Blanc-Féraud was General Program chair of the conference IEEE ISBI 2014 in Beijing.

5.2. Sparse 3D reconstruction in fluorescence imaging

Participants: Emmanuel Soubies, Laure Blanc-Féraud, Sébastien Schaub, Gilles Aubert.

Sparse reconstruction Super-resolution microscopy techniques allow to overstep the diffraction limit of conventional optics. Theses techniques are very promising since they give access to the visualisation of finer structures which is of fundamental importance in biology. In this work we deal with Multiple-Angle Total Internal Reflection Microscopy (MA-TIRFM) which allows reconstructing 3D sub-cellular structures of a single layer of $\sim 300\ nm$ behind the glass coverslip with a high axial resolution. The 3D volume reconstruction from a set of 2D measurements is an ill-posed inverse problem and requires some regularization. Our aim in this work is to propose a new reconstruction method for sparse structures that is robust to Poisson noise and background fluorescence. The sparse property of the solution can be seen as a regularizer using the ℓ_0-norm. Let us denote $\mathrm{f} \in \mathrm{R}^N$ the unknown fluorophore density, then the problem states as

$$\widehat{\mathrm{f}} = \arg\min_{\mathrm{f} \in \mathrm{R}^N} \left(\mathrm{J}_d(\mathrm{f}) + \lambda \|\mathrm{f}\|_0 \right) \tag{14}$$

where J_d is defined from the likelihood function of the observation given f, $\lambda > 0$ is a weight parameter and $\|\cdot\|_0$ denotes the ℓ_0-norm (which counts the number of nonzero components of f). In order to solve this combinatorial problem, we propose a new algorithm based on a smoothed ℓ_0-norm allowing minimizing the non-convex energy (1). Following [20], the idea is to approach the ℓ_0-norm by a suitable continuous function depending on a positive parameter and tending to the ℓ_0-norm when the parameter tends to zero. Then the algorithm solves a sequence of functionals which starts with a convex one (on a large convex set) and introduce progressively the non-convexity of the ℓ_0-norm (Graduated Non Convexity approach). Figure 1 shows the accuracy of the method on a simulated membrane.

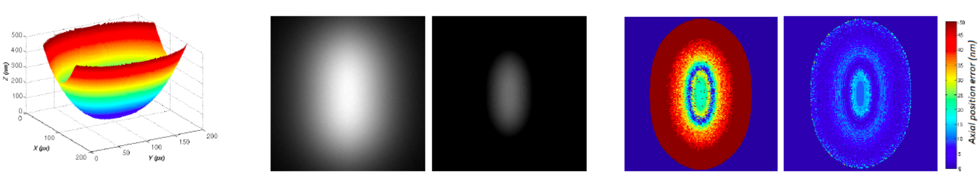


Figure 1. From left to right: Simulated membrane, Microscope acquisition (numerical simulations) with two different incident angles. The two images on the right represent position errors (nm) in the axial direction of the reconstructed membrane obtained with different algorithms: Richardson-Lucy algorithm without regularization (left) and our algorithm with $\lambda = 0.001$ (right).

Axial profile calibration In order to turn on real sample reconstructions we need to perform a calibration of the TIRF microscope. Its principle is based on an evanescent wave with an exponential theoretical decay. However this decay is generally not a pure exponential in practice and we need to have a good knowledge about it. Then based on a phantom specimen of known geometry (bead) we are working on a method to estimate experimentally/numerically this decay profile and calibrate all parameters of the system.

5.3. Penalty analysis for sparse solutions of underdeterminated linear systems of equations

Participants: Emmanuel Soubies, Laure Blanc-Féraud, Gilles Aubert.

In many applications such as compression to reduce data storage, compressed sensing to recover a signal from fewer measurements, source separation, image decomposition and many others, one aims to compute a sparse solution of an underdetermined linear systems of equations. Thus finding such sparse solutions is currently an active research topic. This problem can be formulated as a least squares problem regularized with the ℓ_0-norm. We consider the penalized form

$$\hat{\mathrm{x}} = \arg\min_{\mathrm{x} \in \mathrm{R}^N} \left(\frac{1}{2} \|A\mathrm{x} - d\|^2 + \lambda \|\mathrm{x}\|_0 \right) \tag{15}$$

where $A \in \mathrm{R}^{M \times N}$, $d \in \mathrm{R}^M$ represents the data and $\lambda > 0$ is an hyperparameter characterizing the trade-off between data fidelity and sparsity.

It is well known that reaching a global solution of this $\ell_2 - \ell_0$ functional is a NP-hard combinatorial problem. Besides the non-convexity of this 'norm', its discontinuity at zero makes the minimization of the overall functional a hard task. In this work we focus on non-convex continuous penalties widely used to approximate the ℓ_0-norm which usually lead to better results than the classical ℓ_1 convex relaxation since they are more 'ℓ_0-like'. Based on some results in one dimension, we propose the Exact ℓ_0 penalty (El0). In one dimension and when the matrix A is orthogonal, replacing the ℓ_0-norm in (2) by this penalty gives the convex hull of the overall function. Then we have proved, for any matrix $A \in \mathrm{R}^{M \times N}$, that the global minimizers of the ℓ_2- El0 objective function are the same as for the $\ell_2 - \ell_0$ functional. We also demonstrate that all the local minimizers of this approximated functional are local minimizers for $\ell_2 - \ell_0$ while numerical experiments show that the reciprocal is in general false and that the objective function penalized with El0 admits less local minimizers than the $\ell_2 - \ell_0$ functional. Then, this work provides in some way an equivalence between the initial $\ell_2 - \ell_0$ problem and its approximation using the El0 penalty. One can address problem (2) by replacing the ℓ_0-norm with the El0 penalty which provides better properties for the objective function although the problem remains non-convex. Recently, some authors have proposed algorithms and proved their convergence to critical points of non-smooth non-convex functionals like ℓ_0-El0. Based on such algorithms, we propose a macro algorithm and prove its convergence to a (local) minimizer of the initial $\ell_2 - \ell_0$ functional.

5.4. Motion compensation in two-photon microscopy temporal series

Participants: Caroline Medioni, Grégoire Malandain, Florence Besse, Xavier Descombes.

Acquisitions of 3D image sequences over long period of time, in particular, have enabled neurobiologists to follow complex processes such as the development of neuronal populations or degenerative events occurring in pathological contexts, improving our understanding of the mechanisms involved in brain development and function. In most cases, live samples are moving/growing during long-term imaging. Therefore it is required to compensate for this global 3D motion before measuring the dynamics of the structure of interest. We have proposed a method to compute a coherent 3D motion over a whole temporal sequence of 3D volumes (Figure 2), which is able to capture subtle sub-voxelic displacements.

5.5. Axon Growth Imaging and Modeling

Participants: Agustina Razetti, Caroline Medioni, Florence Besse, Xavier Descombes.

The modeling part of this work has been made in collaboration with S. Komech, E. Pechersky and E. Zhizhina from IITP (Russian Academy of Science)

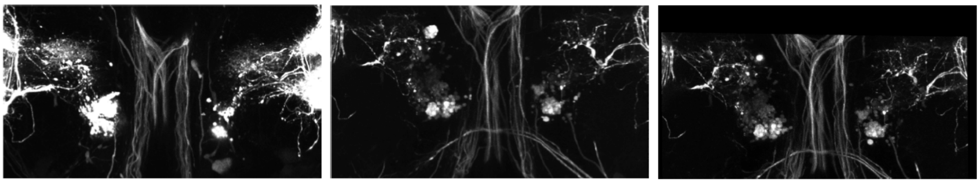


Figure 2. Left and middle: mip views of both the first and the last volumes of a temporal series. Right: mip view of the last volume after motion compensation.

In Drosophila brain, at metamorphosis, Mushroom Body gamma neurons undergo axonal remodeling characterized by a pruning of larval branches followed by regrowth and branching/arborization of adult processes. Axonal regrowth at this stage is essential to consolidate the adult brain and its success is determined by the trajectories followed by the axons and their branches. These trajectories depend on both extracellular guidance signals, and on a complex internal molecular machinery capable to read these signals and act in consequence. F. Besse's team at the IBV Institute has identified genes involved in this regrowth and branching processes [19]. A better understanding of the role of these genes will help to unravel the molecular mechanisms behind these fundamental processes, and lead to a better understanding of the neuronal morphology in both healthy and pathological conditions.

During this PhD project, mathematical and computational tools will be developed to characterize and compare the axonal regrowth and branching dynamics. Different populations of regrowing gamma axons will be considered (i.e. wild type and presenting mutations in relevant genes). The study will be based on both static 3D confocal images of axonal trees, and two-photon in vivo 4D image sequences showing either a single GFP positive regrowing axon or the entire population of regrowing gamma axons marked with GFP. In a first part we have begun to collect data. This includes three parts: i) fly stock maintenance, crossing and selection; ii) sample preparation consisting in pupal brain dissection, medium preparation and sample final assembly; iii) imaging: using different microscopy techniques, eg.confocal/two-photon, microscopy and light sheet microscopy, and acquisition of 4D image sequences. Thanks to an imaging technique developed by C. Medioni in F. Besse's laboratory, we have been able to monitor axonal regrowth and branching at early steps in individual neurons, and to follow them for about 15 hours using the two-photon microscope. These movies will be used in the future to develop the mathematical modeling of axonal regrowth/branching process (see figure 3). Our early works concerning modeling have consisted in investigating some models based on continuous time random walks and characterizing the main axon branch through topological entropy [8] [22].

5.6. Markov Chain for Axon Growth Modeling

Participants: Alejandro Mottini, Xavier Descombes, Florence Besse.

In this work we have defined a 2D discrete stochastic model for the simulation of axonal biogenesis [8]. The model is defined by a third order Markov Chain. The model considers two main processes: the growth process that models the elongation and shape of the neurites and the bifurcation process that models the generation of branches. The growth process depends, among other variables, on the external attraction field generated by a chemoattractant molecule secreted by the target area.

For the validation, we have fluorescently labeled single neurons within intact adult Drosophila fly brains, and have acquired 3D fluorescent confocal microscopy images of their axonal trees. Both normal neurons and neurons in which the function of the imp (mutant type 1) or profilin (mutant type 2) genes was inactivated were imaged. imp encodes a conserved RNA binding protein controlling subcellular mRNA transport and local protein synthesis, and is essential for axonal remodeling. profilin encodes a regulator of the actin cytoskeleton involved in axonal pathfinding. Mutations in these two conserved genes have been linked to neurological pathologies.

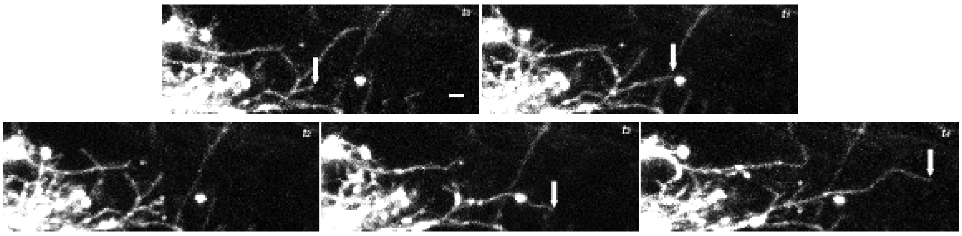


Figure 3. Images extracted from a movie recorded on a two-photon microscope (maximum intensity projection): single axon at the regrowing stage within Drosophila brain. Axons are marked with GFP. t0-t1: elongation step, t2: retraction and branching event, t3: elongation and t4: elongation and branching step. Arrows shows axonal tips and asterisks, the formation of branches. Scale bars for each image: 10μm.

Each image stack has a resolution of $0.093967 \times 0.093967 \times 0.814067 \mu m$ and two channels. The morphology of single axonal trees is visible in the first channel and was manually segmented by an experienced biologist. The morphology of the overall neuronal structure in which axons are developing is visible in the second channel. In total, 53 images (18 normal, 21 type 1 mutant and 14 type 2 mutant) were used. In order to study the attraction field of the populations, all stacks were registered against the first image of the normal population. This was performed using the second channel of each image.

We then have estimated the model parameters to generate two fields for each population, a scalar field that represents the axon flexibility and a vector field that represents the attraction field. Since we obtain some estimates on a sparse set of points in the x, y plane, we extrapolate the fields using a Gaussian Markov Random Field. By qualitatively analyzing the resulting images we have determined that there is no relevant difference on the attraction field between the three populations. We have observed that the field points towards the target area and that its norm is stronger at the starting point of the axons and weaker near the target area, which is consistent with biological expectations. The same procedure was used to analyze the difference on the scalar fields for each population (see Figure 4). In this case we can detect a significant difference between the populations.

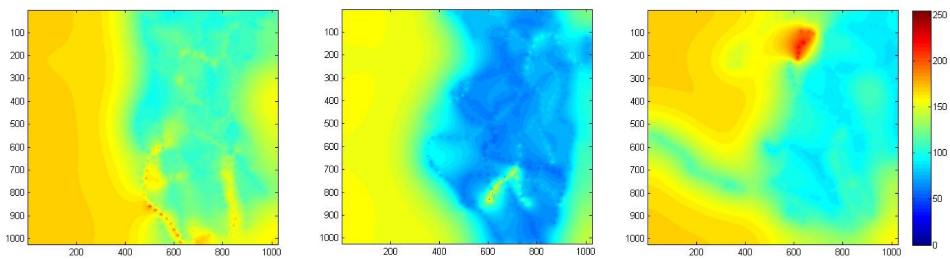


Figure 4. Markov Chain parameter scalar field for the normal (left), mutant type 1 (middle) and mutant type 2 (right) populations.

5.7. Cells detection using segmentation competition

Participants: Emmanuelle Poulain, Emmanuel Soubies, Sylvain Prigent, Xavier Descombes.

Image segmentation has been widely investigated in particular in the context of bioimaging for cells detection. In some cases, the background is clearly identifiable so that a binary mask of the objects can be computed using simple techniques such as thresholding. Therefore, isolated objects are easily recognizable while splitting clusters of objects, which are connected components in the binary mask, remains a challenging task. In fluorescent microscopy devices used for live imaging – *e.g.* confocal, biphoton, Selective Plane Illumination Microscope (SPIM) – an additional difficulty comes from the multiple degradations of the acquired images such as strong noise, spatially varying blur and light attenuation which makes the segmentation a hard task even for selecting a suitable threshold for the background. Since many years, researchers have developed several methods to perform such segmentation. An efficient approach consists in generating seeds that define regions using geometric information through a distance, as in the markers controlled watershed algorithm [21], or image gradient for the active contour approach. These approaches give accurate results providing the seeds are well chosen that is still an open issue. Bayesian approaches, such as marked point process, avoid this bottleneck by selecting randomly generated shapes through the minimization of an energy function. However, they are restricted to low dimensional parametric shapes, such as disks or ellipses, due to computational issues. Tuning the parameters of the segmentation algorithms mentioned above in order to obtain accurate results on the whole image can also be extremely tricky whereas it is much easier to obtain accurate results on different parts of the image using different sets of parameters. To overcome these limits we propose to combine both approaches by generating shapes from state of the art segmentation algorithms using random seeds and/or different sets of parameters. These shapes define a dictionary of candidates from which a competition process, using the Multiple Birth and Cut algorithm, extracts the most relevant shapes. We have validated this selection approach on synthetic data and on a multicellular tumor spheroide slice by comparing the obtained results with two different state of the art segmentation methods to build the dictionary of shapes and compare the performance of our competition approach with the ImageJ particle analyser (see figure 5).

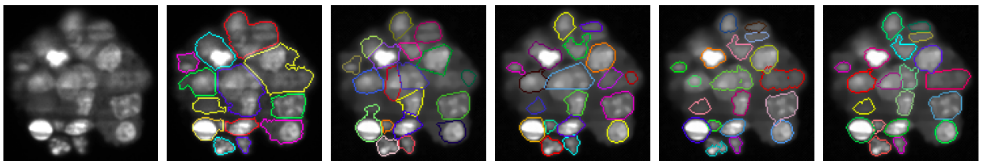


Figure 5. Spheroids of tumor cells stained with a fluorescent nuclear marker (top line left). Fiji Particle Analyzer result (top line right). Results of our approach with dictionary generated by RS-SKIZ with 800 repetitions and a threshold value fixed to 70 (middle line left) MTRS-SKIZ with 80 repetitions and 10 different thresholds (middle line right) MPRS-FM with 100 repetitions and 8 different values for the Fast Marching parameter used to stop the expansion (bottom line left), concatenation of MTRS-SKIZ and MPRS-FM dictionaries (bottom line right).

5.8. Graph cut and attractive interactions

Participants: Tarun Yellamraju, Emmanuel Soubies, Sylvain Prigent, Xavier Descombes.

Marked point processes have proved to be very efficient for segmenting a collection of objects from digital images. The multiple birth and death algorithm provides an optimization framework that allows reasonable time computation. This algorithm has to be embedded in a simluated annealing framework which involves parameters tuning (initial temperature and cooling scheme). This tedious task can be overcame considering a graph cut algorithm instead of the death step. The algorithm then consists in successively adding new random objects in the configuration and selecting the most relevant using the graph cut algorithm. In the graph construction a node is associated to each object. Unfortunately, the regularity condition imposed by the graph cut prevents to consider attractive interactions such as clustering or alignment constraints, which restricts the model to repulsive properties such as non overlap between objects. To overcome this restriction we have investigated new graph constructions by considering nodes defined by clusters of interacting objects. Different strategies have been compared to avoid being tracked in local minima defined by clusters while minimizing the number of required iterations. First results have been obtained on a seeds detection problem (see figure 6).

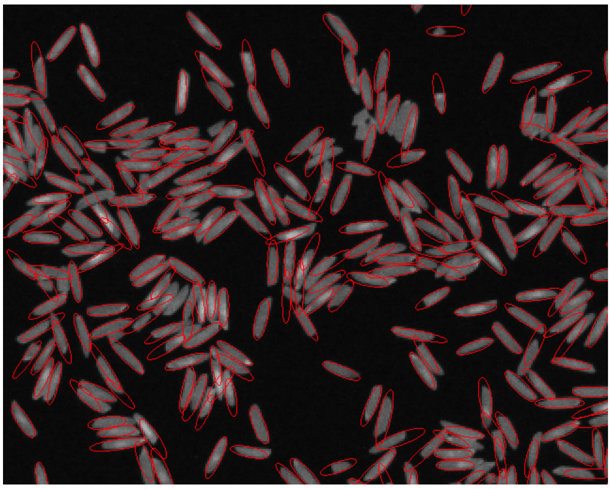


Figure 6. Rice seeds detection using a marked point processes and a birth and cut algorithm (Rice image has been given by Alpha MOS and LAAS-CNRS).

5.9. Cell-to-cell ascidian embryo registration

Participants: Gaël Michelin, Grégoire Malandain.

This work is made in collaboration with Léo Guignard and Christophe Godin (Virtual Plants) and Patrick Lemaire (CRBM), within the Morphogenetics Inria Project Lab.

Recent microscopy techniques allow imaging temporal 3D stacks of developing organs or embryos with a cellular level of resolution and with a sufficient acquisition frequency to accurately track cell lineages. Imaging multiple organs or embryos in different experimental conditions may help to decipher the impact of genetic backgrounds and environmental inputs on the developmental program. For this, we need to precisely compare distinct individuals and to compute population statistics. The first step of this procedure is to develop methods to register individuals.

From a previous work of cell segmentation from microscopy images [6], we propose an approach to extract the Left-Right symmetry plane of embryos at early stages (Figure 7). Then we use the symmetry information to both register these embryos at a similar developmental stage and obtain a cell-to-cell mapping. We assessed the symmetry plane extraction on more than 100 images from 10 individuals between 32-cells and late-neurula development stage. The cell-to-cell registration was performed on 5 distinct individuals at 64-cells and 112-cells stage (Figure 8).

5.10. Quantitative comparison of micro-vasculatures

Participants: Manon Linder, Grégoire Malandain.

This work is made in collaboration with Cécile Duplaa and Thierry Couffinhal (INSERM).

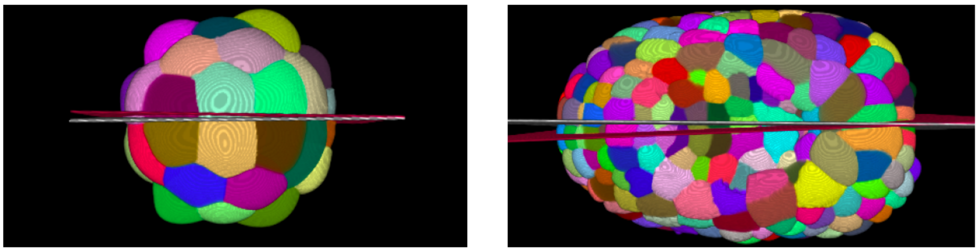


Figure 7. Left-Right symmetry plane initialization (red) and final estimation (white) on (left) a 32-cells stage embryo and (right) a neurula stage embryo.

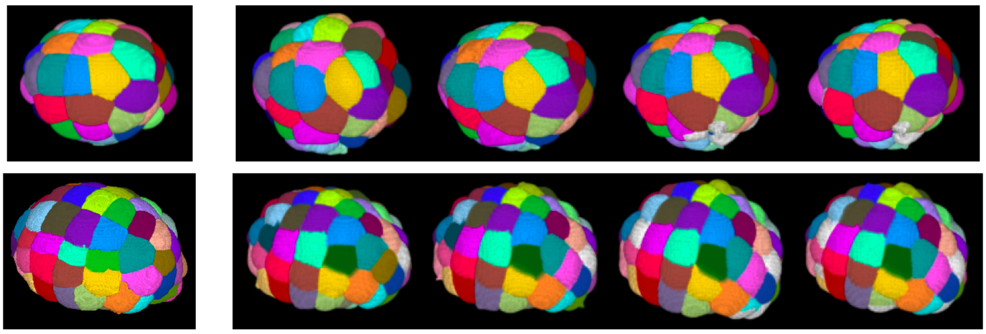


Figure 8. Cell-to-cell mapping between reference image (left) and test images (right) at 64-cells stage (first line) and at 112-cells stage (last line). The reference images are taken from the same individual, the test images are taken at different time-points of a second individual. On the test images, white cells are those that have not been matched to a reference cell.

Angiogenesis is a key component of ontogenesis, but also of tumor development or in some pathology repair (i.e. ischemia). Deciphering the underlying mechanisms of vessel formation is of importance. We aimed at identifying and characterizing the genetic components that are involved in this development. This requires to compare the effect of each gene with respect to the others, hence appeals for quantitative comparisons. We developed a methodology that first transforms a vascular image into a tree and second quantitatively analyze 3D vascular trees (see Figure 9) We conduct real experiments with images of the renal arterial network of different mutant mice, through the development of quantitative measurements that allow for group study.

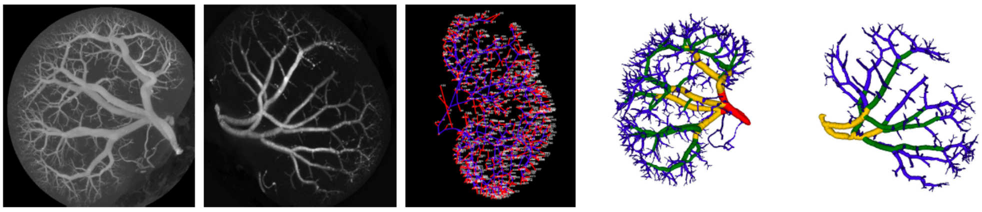


Figure 9. From left to right: two mip views of mouse kidneys acquired with a micro-CT (control and Fzd4 & Fzd6 KO mouse); a labeled tree built from an image; Diameter-defined Strahler classification of arterial trees.

5.11. Pre-clinical molecular dynamic image analysis: 99mTc- pertechnetate biodistribution model of murine stomach with micro-SPECT

Participants: Marine Breuilly, Grégoire Malandain.

This work is jointly conducted with Thierry Pourcher, Jacques Darcourt, Philippe Franken, Kaouthar Chatti, and Philippe Pognonec from the Transporter in Imagery and Oncologic Radiotherapy team (TIRO, CEA-CAL-UNSA).

This project investigates the potential retention of iodide in the stomach, for a better understanding of the iodide biodistribution in the body and more precisely of its potential antiseptic role. To that end, we study the uptake of the ^{99m}Tc-pertechnetate (an iodide analog) within the murine stomach observed thanks to a SPECT camera. Using the coupled SPECT and CT device dedicated to small animals, functional information targeted by a specific radiotracer (^{99m}Tc-pertechnetate) can be imaged dynamically.

The temporal evolution of the uptake is analysed thanks to a dedicated multi-compartment model. The addressed challenges consist in 1) estimating the time-activity curves for the different compartments, and 2) identifying the model parameters.

- ^{99m}-pertechnetate is an iodide analog regarding to the NIS gene. Thus iodide uptake kinetics can be studied through the study of ^{99m}Tc- pertechnetate biodistribution.
- Dynamic SPECT images exhibit a progressive accumulation of ^{99m}Tc-pertechnetate in the stomach wall and diffusion in the stomach cavity.
- The workflow that has been previously proposed in [18] was tested on a larger dataset of five subjects, yielding promising results: The computed model parameters are coherent, and the computed parameter values suggested that there is some iodide retention in the stomach wall.
- A comparison of the dedicated method for extraction of time activity curves with the ones extracted with Pixies software is on-going.
- A comparison of the dedicated method for solving the inverse problem of the compartmental analysis with methods developed by the Turku PET Centre is on-going.

5.12. Massal motility measures to automatically predict fertility scores

Participants: Ana Rita Lopes Simoes, Eric Debreuve.

This work has been done in the scope of the ANR project MOTIMO. We developed a method for automatic scoring of sperm samples in order to predict fertility for the farming industry. The method was applied to samples from rams and goats. A given sample is a video composed of a hundred frames (see Fig. 10).

We analyzed video samples acquired according to four modalities or protocols: drop, chamber, fluorescent beads with a 4x zoom, and fluorescent beads with 10x zoom. Two options have been considered. (1) An optical flow method has been applied to the videos in order to estimate the apparent motion of the seminal fluid (see Fig. 10). Some statistical features of interest (such as entropy) were extracted from the obtained motion fields in order to characterize the sperm massal motility. (2) The second option consisted of tracking the beads on the 4x or the 10x fluorescence videos (see Fig. 10). Some features of interest were also extracted from the resulting bead trajectories. Then using either of these feature sets (obtained with options (1) and (2)), a regression analysis (linear and kernel SVM) was conducted on a subset of the available videos (the learning set) so as to define a prediction function taking features as input and outputting a fertility score. The scores computed by this function were compared to scores assigned by experts. We used the coefficient of determination (commonly denoted by R^2) as a performance measure of the learned prediction function. The best results were obtained using tracking on the ram videos showing fluorescent beads with a 4x zoom ($R^2 = 0.9$). The results on the goat videos were not as satisfying ($R^2 \simeq 0.65$), but the partner providing the videos expressed some concerns about the quality of the acquisition campaign for these data.

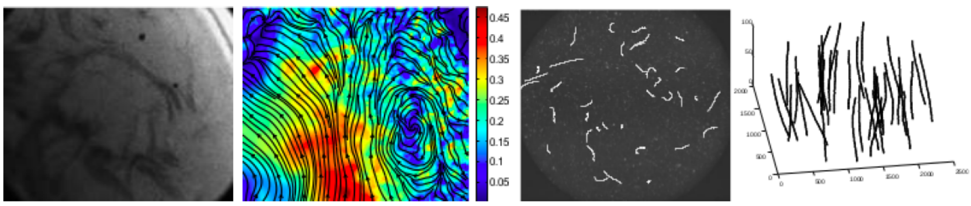


Figure 10. Predicting fertility based on massal motility measures. (Left) One frame of a video for the "drop" modality; (Middle) An example of computed optic flow; (Right: image+plot) An example of bead trajectories.

5.13. Sample selection for SVM learning on large data sets

Participants: Sonia Chaibi, Xavier Descombes, Eric Debreuve.

Support Vector Machines (SVM) represent a popular framework of supervised learning. However, it is not well adapted to large data sets since learning is performed by an optimization procedure involving the whole data set. Yet, in the end, only a small subset of the samples (the so-called support vectors) is retained for prediction. Of course, efficient algorithms exist. Still, it can be interesting to filter out as many samples as possible (the ones that will surely not be part of the support vectors) before initiating the learning procedure.

Sonia Chaibi, a PhD student from UBMA, Algeria, visited the team for a month to collaborate on this subject. The method relies on successive unsupervised sample clustering steps. After each clustering, the homogeneity of the clusters in terms of sample class assignment is used to decide which samples are unlikely to be close to the separation hyperplane (and hence unlikely to be selected as support vectors), and which samples are apparently close to this hyperplane. The former ones can be discarded, thus reducing greatly the number of samples to be processed by the SVM algorithm, while the latter ones are kept, preserving the precision of the separation hyperplane as much as possible.

5.14. Morphological Analysis and Feature Extraction of Neurons from Mouse Cortices Multiscale 3D Microscopic Images

Participants: Alexis Zubiolo, Xavier Descombes, Eric Debreuve.

This work is jointly conducted with Kawssar Harb and Michèle Studer (iBV).

We propose a framework to analyze the morphology of mouse neurons in the layer V of the cortex from 3D microscopic images. We are given 8 sets of images, each of which is composed of a 10x image showing the whole neurons, and a few (2 to 5) 40x images focusing on the somas. The framework consists in segmenting the neurons on both types of images to compute a set of specific morphological features, and in establishing the correspondence between the neurons to combine the features we obtained, in a fully automatic fashion. On the 10x images, we use a multiple birth and cut algorithm to segment the sections of the apical dendrites. Merging these intersections provides the localization of the first branching of the apical dendrite (see Fig. 11 (left)). On the 40x images, we compute an hysteresis threshold to obtain a first segmentation (somas and dendrites starts) and apply iterative morphological operators to reconstruct the full dendrites (see Fig. 11 (middle)). The correspondence map between the two types of images is done using a bipartite graph matching model that associates each neuron configuration of a 40x image – a constellation – to a subset of neurons in the 10 image – the galaxy – (see Fig. 11 (right)).

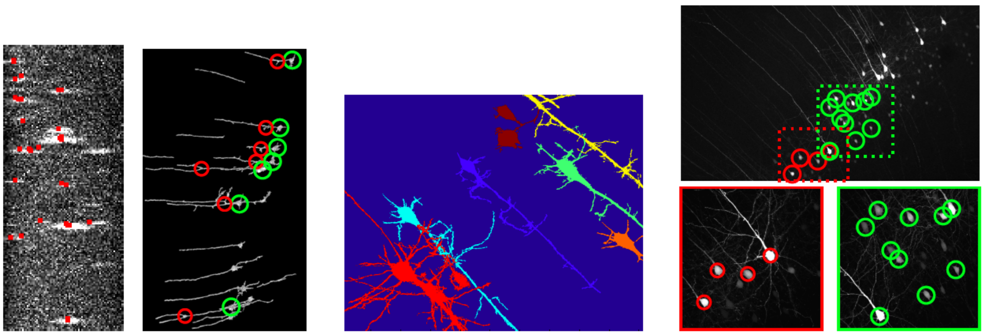


Figure 11. Left: neuron segmentation on the 10x image. Middle: full neuron reconstruction from the 40x image. Right: 10x-40x image maching.

5.15. Whole-Slide Image Analysis of Renal Cell Carcinoma

Participants: Ana Rita Lopes Simoes, Eric Debreuve, Alexis Zubiolo, Xavier Descombes.

This work is jointly conducted with Thierry Pourcher, and Philippe Pognonec (TIRO, CEA-CAL-UNSA), and Damien Ambrosetti (CHU, Nice).

We study histology images of kidney cancer that present different subpopulations of cells (tumor, healthy tissue, stroma, fat, blood, ...). The goal is to analyze the images to help determine the cancer type and stage. Given the resolution of the images ($0.25\mu m$) that leads to very large images (around 100k$\times$100k pixels), a multiscale approach has been considered. At a larger scale, we focus on the cellular architecture and the vascular networks. Regions of interest (ROIs) have been detected with a pixelwise clustering based on neighborhood features (see Fig. 12 (left)). At a smaller scale, we extract more precise information from the cells (nucleus and cytoplasm sizes, shapes and colors, ...). The nuclei of the cells have been segmented using an Hessian determinant-based method (see Fig. 12 (middle)) which enables us to establish statistics about their size. Information on the vascular arborization has been extracted with a Frangi vesselness followed by a cleaning and gap filling post-processing (see Fig. 12 (right)).

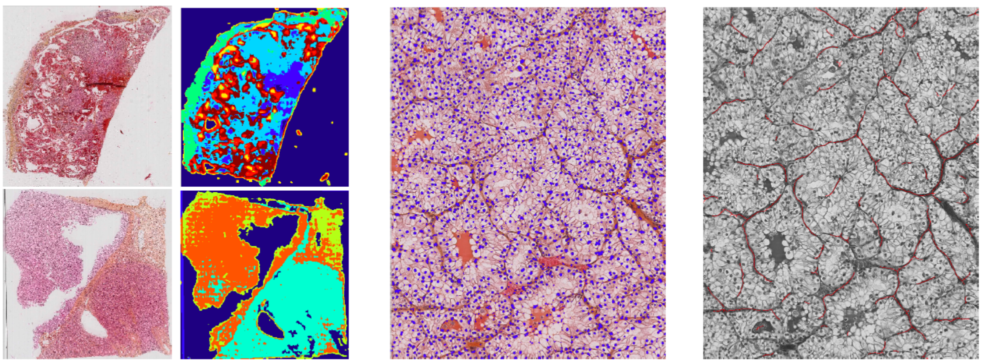


Figure 12. ROI detection (left), Nuclei segmentation in blue (middle) and Vascular arborization extraction in red (right)

MORPHEO Project-Team

6. New Results

6.1. Human Shape and Pose Tracking Using Keyframes

In this work we consider human tracking in multi-view set-ups and investigates a robust strategy that learns online key poses to drive a shape tracking method. The interest arises in realistic dynamic scenes where occlusions or segmentation errors occur. The corrupted observations present missing data and outliers that deteriorate tracking results. We propose to use key poses of the tracked person as multiple reference models. In contrast to many existing approaches that rely on a single reference model, multiple templates represent a larger variability of human poses. They provide therefore better initial hypotheses when tracking with noisy data. Our approach identifies these reference models online as distinctive keyframes during tracking. The most suitable one is then chosen as the reference at each frame. In addition, taking advantage of the proximity between successive frames, an efficient outlier handling technique is proposed to prevent from associating the model to irrelevant outliers. The two strategies are successfully experimented with a surface deformation framework that recovers both the pose and the shape. Evaluations on existing datasets also demonstrate their benefits with respect to the state of the art. This work was presented at CVPR'14 [5].

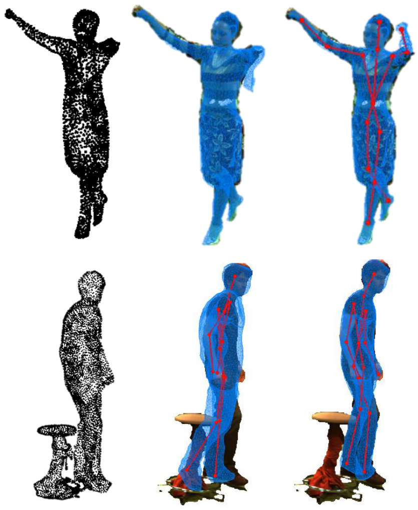


Figure 4. Shape tracking with keframes[5]

6.2. On Mean Pose and Variability of 3D Deformable Models

We present a novel methodology for the analysis of complex object shapes in motion observed by multiple video cameras. In particular, we propose to learn local surface rigidity probabilities (i.e., deformations), and to estimate a mean pose over a temporal sequence. Local deformations can be used for rigidity-based dynamic surface segmentation, while a mean pose can be used as a sequence keyframe or a cluster prototype and has therefore numerous applications, such as motion synthesis or sequential alignment for compression or morphing. We take advantage of recent advances in surface tracking techniques to formulate a generative model of 3D temporal sequences using a probabilistic framework, which conditions shape fitting over all frames to a simple set of intrinsic surface rigidity properties. Surface tracking and rigidity variable estimation can then be formulated as an Expectation-Maximization inference problem and solved by alternatively minimizing two nested fixed point iterations. We show that this framework provides a new fundamental building block for various applications of shape analysis, and achieves comparable tracking performance to state of the art surface tracking techniques on real datasets, even compared to approaches using strong kinematic priors such as rigid skeletons.

Figure 5. Rigidity probability of the shape tracked with mean pose [4]

6.3. Segmentation multi-vues par coupure de graphes

In this paper, we address the problem of object segmentation in multiple views when two or more viewpoints of the same scene are available. We propose a new approach that propagates segmentation coherence information in space, hence allowing evidence in one image to be shared over the complete set. To this aim the segmentation is cast as a single efficient labeling problem over space and time with graph cuts. In contrast to most existing multi-view segmentation methods that rely on some form of dense reconstruction, ours only requires a sparse 3D sampling to propagate information between viewpoints. The approach is thoroughly evaluated on standard multi-view datasets. The obtained results compete with state of the art methods but they are achieved with significantly fewer viewpoints.

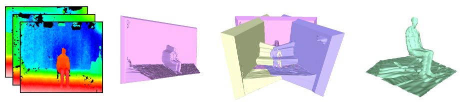


Figure 6. Results of our multi-view segmentation approach over 3 input views, with no user interaction (completely automated). [9]

6.4. Combined Visible and X-Ray 3D Imaging

This work considers 3D imaging of moving objects and introduces a technique that exploits visible and x-ray images to recover dense 3D models. While recent methods such as tomography from cone-beam x-ray can advantageously replace more expensive and higher-dose CT scanners, they still require specific equipment and immobilised patients. We investigate an alternative strategy that combines a single x-ray source and a set of colour cameras to capture rigidly moving samples. The colour cameras allow for coarse marklerless motion tracking, which is further refined with the x-ray information. Once the sample poses are correctly estimated, a dense 3D attenuation model is reconstructed from the set of x-ray frames. Preliminary results on simulated data compared to ground-truth as well as actual in-vivo experiments were presented at the conference MIUA'14 [6].

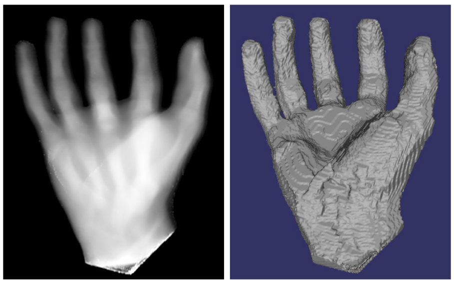


Figure 7. Human hand reconstruction with a combined visible and x-ray system [6]

6.5. Non-Rigid Registration meets Surface Reconstruction

Non rigid registration is an important task in computer vision with many applications in shape and motion modeling. A fundamental step of the registration is the data association between the source and the target sets. Such association proves difficult in practice, due to the discrete nature of the information and its corruption by various types of noise, e.g. outliers and missing data. In this work we investigate the benefit of the implicit representations for the non-rigid registration of 3D point clouds. First, the target points are described with small quadratic patches that are blended through partition of unity weighting. Then, the discrete association between the source and the target can be replaced by a continuous distance field induced by the interface. By combining this distance field with a proper deformation term, the registration energy can be expressed in a linear least square form that is easy and fast to solve. This significantly eases the registration by avoiding direct association between points. Moreover, a hierarchical approach can be easily implemented by employing coarse-to-fine representations. Experimental results were conducted with point clouds from multi-view data sets. The qualitative and quantitative comparisons show the outperformance and robustness of our framework. This work was presented at 3DV'14[7].

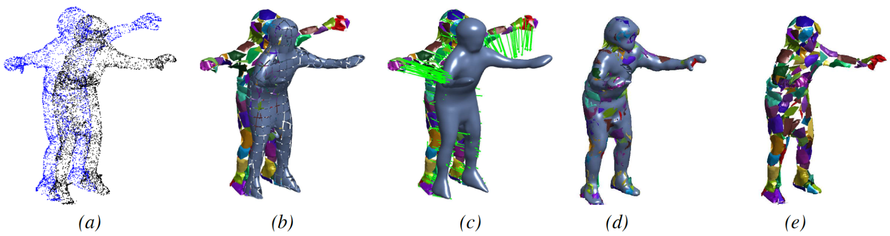


Figure 8. Using implicit interface for registration: (a) initial pose of the source and target sets; (b) source patches and the local quadrics representing the target; (c) the implicit interface induces a gradient field; (d) deformed source patches fitting the interface; a coarse-to-fine interface has been used in (c) - (d); (e) the final deformation of the template [7].

6.6. High Resolution 3D Shape Texture from Multiple Videos

We examine the problem of retrieving high resolution textures of objects observed in multiple videos under small object deformations. In the monocular case, the data redundancy necessary to reconstruct a high-resolution image stems from temporal accumulation. This has been vastly explored and is known as super-resolution. On the other hand, a handful of methods have considered the texture of a static 3D object observed from several cameras, where the data redundancy is obtained through the different viewpoints. We introduce a unified framework to leverage both possibilities for the estimation of a high resolution texture of an object. This framework uniformly deals with any related geometric variability introduced by the acquisition chain or by the evolution over time. To this goal we use 2D warps for all viewpoints and all temporal frames and a linear projection model from texture to image space. Despite its simplicity, the method is able to successfully handle different views over space and time. As shown experimentally, it demonstrates the interest of temporal information that improves the texture quality. Additionally, we also show that our method outperforms state of the art multi-view super-resolution methods that exist for the static case. This work was presented at CPVR'14 [8].

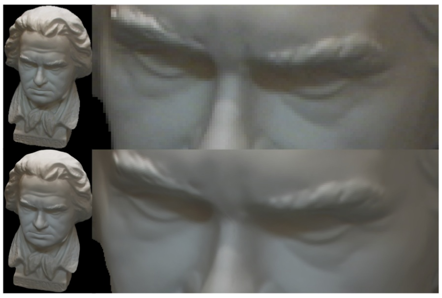


Figure 9. Input view 768 × 576 resolution with up-sampling by factor of three, BEETHOVEN dataset. Super-resolved 2304×1728 output of our algorithm rendered from identical viewpoint [8].

MULTISPEECH Team

6. New Results

6.1. Highlights of the Year

The version 2 of our source separation toolbox FASST [65] has been downloaded more than 300 times since its release in January 2014.

6.2. Explicit modeling of speech production and perception

Participants: Yves Laprie, Slim Ouni, Vincent Colotte, Anne Bonneau, Agnès Piquard-Kipffer, Martine Cadot [Univ. Lorraine], Antoine Liutkus, Emmanuel Vincent, Odile Mella, Benjamin Elie, Camille Fauth, Julie Busset, Andrea Bandini, Guillaume Gris, Simon Meoni.

6.2.1. Articulatory modeling

6.2.1.1. Acquisition of articulatory data

Acquisition of articulatory data plays a central role in the construction of articulatory models and investigation of articulatory gestures. In cooperation with the IADI laboratory (Nancy hospital) we thus conducted a series of preliminary experiments intended to acquire cine-MRI data. Images of the film are reconstructed thanks to the cine-GRICS algorithm developed at IADI [56].

The second research track concerns ultrasound (US) imaging which presents the interest of offering a good temporal resolution without any health hazard and at a reasonable price. However, it cannot be used alone because there is no reference coordinate system and no spatial calibration. We thus used a multimodal acquisition system developed by the Magrit team, which uses electromagnetography sensors to locate the US probe, and the method used to calibrate the US modality. We experimented this system to investigate the most appropriate acquisition protocol for Magnetic Resonance Imaging [37].

We also use an articulograph to acquire articulatory data. Within the framework of the EQUIPEX ORTOLANG, we acquired this year an AG501, a 24-channel articulograph. This system is the most advanced electromagnetography acquisition system. It has been used for two articulatory studies: (1) investigating the effects of posture and noise on speech production [48] and (2) studying the pauses in spontaneous speech from an articulatory point of view. We also conducted an exploratory study on retrieving the 3D shape of the palate from electromagnetography tracings (the work of Simon Meoni, a master student in Cognitive Sciences).

6.2.1.2. Acoustic-to-articulatory inversion

Our previous works about acoustic-to-articulatory inversion relied on the exploration of a vast articulatory codebook covering the whole articulatory space that could be reached by a speaker. This solution presents the main drawback of requiring the construction a codebook for each speaker. We thus developed a multimodal approach to estimate the area function and the length of the vocal tract of oral vowels. The method is based on an iterative technique consisting in deforming an initial area function so that the output acoustic vector matches a specified target. The chosen acoustic vector is the formant frequency pattern. In order to regularize the ill-defined problem, several constraints are added to the algorithm. First, the lip termination area is estimated via a facial capture software. Then, the area function is constrained so that it does not get too far from a neutral position, and so that it does not change too quickly from a temporal frame to the next, when dealing with dynamic inversion. The method proves to be efficient for approximating the area function and the length of the vocal tract for oral French vowels, both in static and dynamic configurations.

6.2.1.3. Articulatory models

The development of articulatory models is a crucial aspect of articulatory synthesis since this determines the success of synthesis. The previous model was developed for X-ray images. This means that the laryngeal part of the model associates the larynx with the piriform sinuses event if these two structures are not in the same sagittal plane. The new model separates the two structures if needed. Additionally, the larynx and the epiglottis are controlled independently which corresponds to the anatomical truth. Previous attempts to modeling epiglottis used principal component analysis applied to the contours drawn on X-ray images. Unfortunately the width of the epiglottis varies from one image to the other and PCA thus learns a spurious "inflating" component. The new model uses the epiglottis centerline plus a constant width which prevents this error.

The second major improvement concerns the use of virtual targets in the construction of the articulatory model. Virtual targets are used to separate the contribution of the tongue contour from those of the palate. The objective it to render the articulation of consonants more correctly since they require a contact between the tongue and the palate at a very precise point [38].

These two improvements of the articulatory model were used in the articulatory copy synthesis experiments [11].

The construction of models was also tackled from a data mining point of view. A robust data mining approach was designed to automatically extract complex statistically significant connections between data (e.g. interactions between more than two variables). This work could be used for data other than X-ray images [54].

6.2.2. Expressive acoustic-visual synthesis

Right now, we are investigating the state-of-the-art of the field of expressive speech and how to acquire efficiently expressive speech corpus. As a first step, we are also investigating visual acquisition techniques to track facial expression. This is the work of the visiting PhD student Andrea Bandini (from University of Bologna). Another step toward expressive speech synthesis is to have an expressive face model. In this context, the expressivity is mainly based on the dynamics. In fact, when the human facial movements are natural and accurately replicated on the 3D model, we can reach a reasonable expressivity. In this context, we are conducting new research toward an expressive talking head. In this context, we acquired a high-resolution 3D model of a human speaker and we are developing methods to animate the model using motion capture data. This was the work the master student Guillaume Gris. We also investigated the advantage of generating visual speech from sequences of 2D Images, when the 3D data is lacking [43].

6.2.3. Categorization of sounds and prosody for native and non-native speech

Categorization of sounds and prosody for non-native speech is the object of the ANR+DFG project IFCASL devoted to French and German languages. Within this project, we built a bilingual corpus and started a study about the realization of (final) voicing in both languages. We also gave a training course about non-native phonetic realizations for a Spring School devoted to *Individualized centered approaches to speech processing* [63].

6.2.3.1. Bilingual speech corpus of French and German language learners

We designed a corpus of native and non-native speech for the French-German language pair, with a special emphasis on phonetic and prosodic aspects. To our knowledge there is no suitable corpus, in terms of size and coverage, currently available for this target language pair [9].

We adopted a two step process to create the corpus. Firstly, a bilingual corpus including all sounds of each language and all speech phenomena of potential interest was recorded from a few speakers (14), and analyzed. Its analysis revealed/confirmed: 1) the existence of special strategies due to sentence reading and sentence listening conditions, 2) the importance of recording duration (the recording sessions should not last more than one hour to avoid subjects' fatigue), 3) the frequence and importance of some mispronunciations (voicing problems, erroneous presence (or absence) of /h/ for German (or French) non-native speakers, rhythm ...). Secondly, we specified and collected the final corpus [24], which is focused on the problems revealed by the preliminary corpus. One hundred speakers (50 French and 50 German speakers), beginners and advanced speakers, recorded 60 sentences in their second language and 30 in their native language, which gave a total amount of about 6000 non-native and 3000 native sentence realizations. The sentences were read in two conditions depending upon whether or not the subjects listen to a reference before producing the sentence. A small text as well as sentences devoted to focus analysis completed the corpus. The data was segmented and labelled at word and phone levels by an automatic alignment algorithm elaborated by our team (cf. 6.4.3.2). The outputs were then manually checked at the levels of phones and words (phonetic transcription) and corrections were made if necessary. In order to check the homogeneity of the corrections made by the seven annotators, phone boundaries were compared with those achieved by a golden annotator on a few sentences using the CoALT tool.

6.2.3.2. Devoicing of final obstruents by German learners

We investigated a typical example of L1-L2 interference: the realization of voiced fricatives in final position, where the opposition between voiced and unvoiced consonants is neutralized in German (with a bias towards unvoiced consonants) but not in French. As a consequence, German speakers learning French as a second language often produce unvoiced fricatives in final position instead of the expected voiced consonants. We analysed the production of French voiced fricatives for 40 non-native (beginners and advanced speakers) and 8 native speakers. We measured the ratio of locally unvoiced frames in the consonantal segment and also the ratio of consonantal duration vs. the duration of the preceding vowel. Results showed that the realizations of French fricatives by German speakers varied with speakers, speakers' level and experimental condition (there were two conditions depending on whether or not the subjects listened to a reference before producing the sentence) [23]. As could be expected we observed a continuum between typically voiced and typically unvoiced realizations, and best level speakers tend to produce more typically French realizations. Our next study will concern the perceptual identification of learners' realizations and the link between perceptual answers and acoustic cues values.

6.3. Complex statistical modeling of speech

Participants: Emmanuel Vincent, Antoine Liutkus, Denis Jouvet, Dominique Fohr, Irina Illina, Joseph Di Martino, Emad Girgis, Arseniy Gorin, Nathan Souviraà-Labastie, Luiza Orosanu, Imran Sheikh, Xabier Jaureguiberry, Baldwin Dumortier.

6.3.1. Acoustic modeling

6.3.1.1. Theory for audio source separation

Our work on audio source separation was marked by the release of version 2 of our toolbox FASST, which was demonstrated at ICASSP 2014 [65], and by the publication of a review paper about guided audio source separation for *IEEE Signal Processing Magazine* [16]. Audio source separation is an inverse problem, which requires the user to guide the separation process using prior models for the source signals and the mixing filters or for the source spectra and their spatial covariance matrices.

On the topic of the mixing parameters, we studied the impact of sparsity penalties over the mixing filters [8] and deterministic subspace constraints [10] over the spatial covariance matrices.

Modelling the spectra of the sources is a fundamental problem in source separation, that aims at catching their main features while requiring few parameters to estimate. We proposed a new framework called Kernel Additive Modelling (KAM). In contrast to Nonnegative Matrix Factorization approaches (NMF), KAM permits to model sources spectro-temporal evolutions only locally. It generalizes many methods from the state-of-the-art, including REPET (voice/music separation) and HPSS (harmonic/percussive separation) and is the first framework to settle them on principled statistical grounds. This year, we have thus been very active not only in diffusing REPET and its variants to a large audience, notably through the publication of a chapter book on the topic [58], but also by establishing many international collaborations on KAM, leading to the publication of one journal paper in IEEE TSP [13] and to two international conference papers [25], [42].

In parallel, we started a new research track on the fusion of multiple source separation techniques. In the specific case when the source spectra are modeled by NMF, the number of components of the NMF is known to have a noticeable influence on separation quality. Many methods have been proposed to select the best order for a given task. To go further, we proposed to use model averaging. As existing techniques do not allow an effective averaging, we introduced a generative model in which the number of components is a random variable and we proposed a modification to conventional variational Bayesian (VB) inference. Initial experiments showed promising results [33], [32].

6.3.1.2. Audio separation based on multiple observations

An interesting scenario for informed audio source separation is when the signals to separate can be observed through deformed references. We proposed a general approach for the separation of multichannel mixtures guided by multiple, deformed reference signals such as repeated excerpts of the same music or repeated versions of the same sentence uttered by different speakers [46], [66].

A related topic is the removal of interferences from live recordings. In this scenario, there are as many microphones as source signals, but each microphone captures not only its dedicated source, but also some interference from the other ones. We proposed a variant of KAM, called KAM for Interference Removal (KAMIR) that permits to address this scenario. The corresponding study has been achieved in collaboration with New York and Erlangen universities.

6.3.1.3. Separation and dereverberation

In order to complement source separation by dereveberation of the source signals, we devoted some work to the estimation of the reverberation time (RT60). In many situations, the room impulse response (RIR) is not available and the RT60 must be blindly estimated from a speech or music signal. Current methods often implicitly assume that reverberation dominates direct sound, which restricts their applicability to relatively small rooms or distant sound sources. We proposed a blind RT60 estimation method that is independent of the room size and the source distance and showed that the estimation error is significantly reduced even in the case when reverberation dominates [21].

6.3.1.4. Corpora for audio separation

Finally, we pursued our long-lasting efforts on the evaluation of audio source separation by providing more details about the DEMAND dataset, that is the first-ever publicly available dataset of multichannel real-world noise recordings [55]. Furthermore, we have continued our efforts on providing corpora for the evaluation of music source separation methods (notably for music/voice separation) and target at significantly extending the SiSEC corpus in 2015 to several hundreds complete recordings, to be used for the first time at SiSEC 2015.

6.3.1.5. Detailed acoustic modeling

Acoustic models aim at representing the acoustic features that are observed for the sounds of the language, as well as for non-speech events (silence, noise, ...). Currently context-dependent hidden Markov models (CD-HMM) constitute the state of the art for speech recognition. However, for text-speech alignment, simpler context-independent models are used as they provide better performance.

In conventional HMM-based approaches that rely on Gaussian mixture densities (GMM), the Gaussian components are estimated independently for each density. Thus, we have focused recent studies on enriching the acoustic models themselves in view of handling trajectory and speaker consistency in decoding. A new modeling approach was developed that takes advantage of the multiple modeling ideas and involves a sharing of parameters. The idea is to use the multiple modeling approach to partition the acoustic space according to classes (manual classes or automatic classification). Then, for each density, Gaussian components are estimated using the data associated to the classes. These class-based Gaussian components are then pooled to provide the set of Gaussian components of the density. Finally class dependent mixture weights are estimated for each density; such approach allows us to better parameterize GMM-HMM without increasing significantly the number of model parameters. Experiments on French radio broadcast news data demonstrated the improvement of the accuracy with such parameterization compared to models with a similar, or even a larger number of parameters. Another approach has been proposed that combines the structuring of the Gaussian components of the densities with respect to some data classes, with the stranded-based approach which introduces probabilities for the transitions between the Gaussian components of the densities when moving from one frame to the next. A detailed analysis of stranded GMM was conducted on data containing different types of non-phonetic variability [29]. The combination of stranded GMM with class-structured densities was evaluated on an English connected digits task using adult and child data [27] and for phonetic decoding on a larger French telephone speech database [26]. This approach was later combined with feature normalization [28].

6.3.1.6. Robust acoustic modeling

In the framework of using speech recognition for helping communication with deaf or hard of hearing people, robustness of the acoustic modeling is investigated. Current studies relate to improving robustness with respect to speech signal level and environment noise through multicondition training and enhanced set of acoustic features.

6.3.1.7. Unsupervised acoustic model training

In previous experiments relating to the combination of speech decoder outputs for improving speech recognition performance [4], it was observed that when a forward-based and a backward-based decoder were providing a same word hypothesis, such common word hypothesis is correct in more than 90% of the cases [71]. Hence, we have investigated how such behavior can help for selecting data for unsupervised training of acoustic models. Best performance is achieved when selecting automatically transcribed data (speech segments) that have the same word hypotheses when processed by the Sphinx forward-based and the Julius backward-based transcription systems, and this selection process outperforms confidence measure based selection. Overall, selecting automatically transcribed speech segments that have the same word hypotheses for the two speech transcription systems, and adding this automatically transcribed and selected data to the manually transcribed data leads to significant word error rate reductions on the ESTER2 data (radio broadcast news) when compared to the baseline system trained only on manually transcribed speech data [34].

6.3.1.8. Score normalization

Existing techniques for robust ASR typically compensate distortion on the features or on the model parameters themselves. By contrast, a number of normalization techniques have been defined in the field of speaker verification that operate on the resulting log-likelihood scores. We provided a theoretical motivation for likelihood normalization due to the so-called "hubness" phenomenon and we evaluated the benefit of several normalization techniques on ASR accuracy for the 2nd CHiME Challenge task. We showed that symmetric normalization (S-norm) reduces the relative error rate by 43% alone and by 10% after feature and model compensation [53].

6.3.2. Linguistic modeling

6.3.2.1. Out-of-vocabulary proper name retrieval

Recognition of proper names is a challenging task in information retrieval in large audio/video databases. Proper names are semantically rich and are usually key to understanding the information contained in a document. Within the ContNomina project (cf. 8.1.4), we focus on increasing the vocabulary coverage of a speech transcription system by automatically retrieving proper names from contemporary text documents. We proposed methods that dynamically augment the automatic speech recognition system vocabulary, using lexical and temporal features in diachronic documents (documents that evolve over the time). Our work uses temporal context modeling to capture the lexical information surrounding proper names so as to retrieve out-of-vocabulary proper names and increase the ASR vocabulary size. We focus on exploiting the lexical context based on temporal information from diachronic documents. Our assumption is that time is an important feature for capturing name-to-context dependencies. We also studied different metrics for proper name selection in order to limit the vocabulary augmentation: a method based on Mutual Information and a new method based on cosine-similarity measure. Recognition results show a significant reduction of the proper name error rate using augmented vocabulary [30][31].

6.3.2.2. Hybrid language modeling

In the framework of using speech recognition for helping communication with deaf or hard of hearing people, the handling of out-of-vocabulary words is a critical aspect. Indeed, the size of the vocabulary is always limited (even if large or very large), and the system is not able to recognize words out of its lexicon. Such words would then be transcribed as sequences of short words which involve similar sounds as the unknown word. However the interpretation of such sequences of small word require a lot of efforts. Hence the idea of combining in a single model a set of words (the most frequent and/or most relevant for the application context) and a set of syllables. With such an approach,unknown words are usually recognized as sequences of syllables which are easier to interpret. By setting different thresholds on the confidence measures associated to the recognized words (or syllables), the most reliable word hypotheses can be identified, and they have correct recognition rates between 70% and 92% [44][45].

6.3.2.3. Music language modeling

Similarly to speech, music involves several levels of information, from the acoustic signal up to cognitive quantities such as composer style or key, through mid-level quantities such as a musical score or a sequence of chords. The dependencies between mid-level and lower- or higher-level information can be represented through acoustic models and language models, respectively. We pursued our pioneering work on music language modeling, with a particular focus on the modeling of long-term structure [20]. We also proposed a new Bayesian n-gram topic modeling and estimation technique, which we applied to genre-dependent modeling of chord sequences and to music genre classification [15].

6.3.3. Speech generation by statistical methods

6.3.3.1. Enhancing pathological voice by voice conversion techniques

Enhancing the pathological voice in order to make it more intelligible would allow persons having this kind of voice to communicate more easily with those around them. In our group we chose to improve the pathological voice by means of voice conversion techniques. Since we began this study, we have succeeded to predict the complete magnitude spectrum. In doing so, we free ourselves from the prediction of the fundamental frequency of speech (F0). Such an interesting result allows us to obtain converted speech of good audio quality. Now in order to obtain perfect conversion, we are trying, with Emad Girgis, a postdoctoral student who began his work in November 2014, to predict the phase spectrum. To achieve this goal, Emad intends to use Deep Neural Networks (DNN). We expect first results in the beginning of 2015.

6.3.3.2. Enhancing pathological voice by voice recognition techniques

Another possibility for enhancing the pathological voice is to recognize it. Othman Lachhab, a PhD student, is working on the recognition of the esophageal voice: using high order temporal derivatives combined with an Heteroscedastic Linear Discriminant Analysis (HLDA) he reached an interesting phone recognition rate of 63.59% [36]. Currently Othman, is trying to improve his results by using voice conversion techniques. Using these techniques pathological features are projected in a clean-natural speech feature space, and preliminary results exhibit an increase of 1.70% of the phone recognition rate.

6.3.3.3. F0 detection using wavelet transforms

Another possible interesting track for improving voice conversion techniques is to predict the fundamental frequency of speech. For doing so, it is necessary to have a good F0 detector. As part of her thesis, Fadoua Bahja developed many F0 detection algorithms [69] [1]. The latest, using wavelet transform for denoising the cepstrum signal, has been submitted for publication in an international journal.

6.4. Uncertainty estimation and exploitation in speech processing

Participants: Emmanuel Vincent, Dominique Fohr, Odile Mella, Denis Jouvet, Agnès Piquard-Kipffer, Dung Tran.

6.4.1. Uncertainty and acoustic modeling

In many real-world conditions, the speech signal is overlapped with noise, including environmental sounds, music, or undesired extra speech. Speech enhancement is useful but insufficient: some distortion remains in the enhanced signal which must be quantified in order not to be propagated to the subsequent feature extraction and decoding stages. The framework of uncertainty decoding assumes that this distortion has a Gaussian distribution and seeks to estimate its covariance matrix [5]. A number of uncertainty estimators and propagators have been proposed for this purpose, which typically operate on diagonal covariance matrices and are based on fixed mathematical approximations or heuristics. We obtained more accurate uncertainty estimates by propagating the full uncertainty covariance matrix and by fusing multiple uncertainty estimators [50], [51]. Overall, we obtained 18% relative error rate reduction with respect to conventional decoding (without uncertainty), that is about twice as much as the reduction achieved by the best single uncertainty estimator and propagator.

In order to motivate further work by the community, we created a new international evaluation campaign on that topic in 2011: the CHiME Speech Separation and Recognition Challenge [2]. After two successful editions in 2011 and 2013, we started working and collecting a new corpus towards the organization of a third edition to be announced in 2015.

6.4.2. Uncertainty and speech recognition

In the framework of using speech recognition for helping communication with deaf or hard of hearing people in the FUI project Rapsodie (cf. 8.1.5), our goal is to find the best way for displaying the speech transcription results. To our knowledge there is no suitable, validated and currently available display of the output of automatic speech recognizer for hard-of-hearing persons, in terms of size, colors and choice of the written symbols. The difficulty comes from the fact that speech transcription results contain recognition errors, which may impact the understanding process. Although the speech recognition system does not know the errors it makes, through the computation of confidence measures, the speech recognizer estimates if a word or a syllable is rather correctly recognized or not (cf. 6.3.2.2); hence such information can be used to adjust the display of the transcription results.

We have adopted a two-step process. Firstly, we conducted a feasibility study with three hard-of-hearing persons including written display tests on print media and interviews. Secondly, we set up an experimental protocol with five hard-of-hearing persons. It included comprehension tests of 40 written sentences recorded by a French native speaker video projected onto a screen. We have also conducted parallel interviews. Their analysis revealed: (1) the interest of the participants in the project; (2) their difficulties to read International Phonetic Alphabet; (3) the importance of knowing the context of communication; (4) the need for aid in case of errors of the speech recognition system by emphasing the words that are supposed to be well recognized by the system. At this stage of the experimental period, the best display associates writing in a bold spelling the words that are supposed to be correctly recognized, and writing in a normal font using simplified French phonetics the words that are possibly wrongly recognized (according to their confidence measure). The next step will be to set up another experimental protocol in order to compare the current display in three conditions (written sentences vs written sentences with oral and lip reading vs lip reading only).

6.4.3. Uncertainty and phonetic segmentation

As described below, phonetic segmentation has been studied this year for spontaneous speech and non-native speech. Moreover, some portions (of about 30 secondes) of various speech documents have been manually annotated (checking and correction of an automatic segmentation). In the future this manually annotated data will be used to analyze the accuracy of the automatic segmentation, and also to elaborate measures that estimate the quality of the segmentation.

6.4.3.1. Alignment with spontaneous speech

Within the ANR ORFEO project (cf. 8.1.2), we addressed the problem of the alignment of spontaneous speech. The ORFEO audio files were recorded under various conditions with a large SNR range and contain extra speech phenomena and overlapping speech. We trained several sets of acoustic models and tested different methods to adapt them to the various audio files. For selecting the best acoustic models, we compared the alignment outputs obtained with the different acoustic models by using our tool CoALT and the manually annotated portions described above.

We also designed a new automatic grapheme-phoneme tool to generate the potential pronunciations of words and proper names. For what concerns overlapping speech, among the different orthographic transcripts corresponding to the overlapping area, we determined as the main transcript the one that best matches the audio signal, the others are kept in other tiers (in a Praat TextGrid file) with the same time boundaries.

6.4.3.2. Alignment with non-native speech

Non-native speech alignment with text is one critical step in computer assisted foreign language learning [3]. The alignment is necessary to analyze the learner's utterance, in view of providing some prosody feedback (as for example bad duration of some syllables). However, non-native speech alignment with text is much more complicated than native speech alignment. This is due to the pronunciation deviations observed on non-native speech, as for example the replacement of some target language phonemes by phonemes of the mother tongue, as well as errors in the pronunciations. Non-native speech alignment with text is currently studied in the ANR IFCASL project (see 8.1.3).

6.4.4. Uncertainty and prosody

A statistical analysis was conducted on a large annotated speech corpus to investigate the links between punctuation and automatically detected prosodic structures. The speech data comes froms radio broadcast news and TV shows, that were manually annotated during French speech transcription evaluation campaigns. These corpora contain more than 3 million words and almost 350,000 punctuation marks. The detection of the prosodic boundaries and of the prosodic structures is based on an automatic approach that integrates little linguistic knowledge and mainly uses the amplitude and the direction of the F0 slopes, as well as phone durations. A first analysis of the occurrences of the punctuation marks, with respect to various sub-corpora, has highlighted the variability among annotators. Then, a detailed analysis of the prosodic parameters with respect to the punctuation marks, whether followed or not by a pause, and of the links between the automatically detected prosodic structures and the manually annotated punctuation marks was conducted [18].

MUSE Team

6. New Results

6.1. Pinpointing Home and Access Network Delays Using WiFi Neighbors

Participants: Lucas Di Cioccio (LIP6/Technicolor), Martin May (Technicolor), Jim Kurose (University of Massachusetts, Amherst), Renata Teixeira

Home Internet users and Internet access providers need tools to assist them in diagnosing and troubleshooting network performance problems. Today, expert users may rely on simple techniques using round-trip-measurements to local and remote points to locate delays on an end-to-end path. Unfortunately, round-trip measurements do not provide accurate diagnoses in the presence of asymmetric link capacities and performance, which is often the case in residential access. Our work [8] introduces *neighbor-assisted delay diagnosis* (NADD) - an approach for pinpointing the location of delays (among the home, access, and wide-area network), leveraging end-host multi-homing capabilities. NADD runs on an end host connected simultaneously to the home gateway and to a neighbor WiFi access point. Our evaluation shows that NADD efficiently detect and distinguish uplink and downlink delays with small error. In addition, we learn from a proof-of-concept deployment in five homes in France that our techniques can work "in the wild." Technicolor filed a patent on this work [8].

6.2. Locating Throughput Bottlenecks in Home Networks

Participants: Srikanth Sundaresan (ICSI), Nick Feamster (Princeton), Renata Teixeira

We developed *WTF (Where's The Fault?)* [4], a system that localizes performance problems in home and access networks. We implement WTF as custom firmware that runs in an off-the-shelf home router. WTF uses timing and buffering information from passively monitored traffic at home routers to detect both access link and wireless network bottlenecks. The Federal Communication Commission (FCC) in the United States deployed WTF in 3000 homes for a few days in November 2014. We are currently analyzing the resulting dataset to help shed light on common pathologies that occur in home networks.

6.3. Measuring the Performance of User Traffic in Home Wireless Networks

Participants: Srikanth Sundaresan (ICSI), Nick Feamster (Princeton), Renata Teixeira

This work [5] studies how home wireless performance characteristics affect the performance of user traffic in real homes. Previous studies have focused either on wireless metrics exclusively, without connection to the performance of user traffic; or on the performance of the home network at higher layers. In contrast, we deploy a passive measurement tool on commodity access points to correlate wireless performance metrics with TCP performance of user traffic. We implement our measurement tool, deploy it on commodity routers in 66 homes for one month, and study the relationship between wireless metrics and TCP performance of user traffic. We find that, most of the time, TCP flows from devices in the home achieve only a small fraction of available access link throughput; as the throughput of user traffic approaches the access link throughput, the characteristics of the home wireless network more directly affect performance. We also find that the 5 GHz band offers users better performance better than the 2.4 GHz band, and although the performance of devices varies within the same home, many homes do not have multiple devices sending high traffic volumes, implying that certain types of wireless contention may be uncommon in practice.

6.4. Characterizing Bufferbloat and its Impact at End-hosts

Participants: Stephane Wustner, Jaideep Chandrashekar (Technicolor), Renata Teixeira

While, on routers and gateways, buffers on forwarding devices are required to handle bursty Internet traffic, overly large or badly sized buffers can interact with TCP in undesirable ways. This phenomenon is well understood and is often called “bufferbloat”. Although a number of previous studies have shown that buffering (particularly, in home) can delay packets by as much as a few seconds in the worst case, there is less empirical evidence of tangible impacts on end-users. In [3], we develop a modified algorithm that can detect bufferbloat at individual end-hosts based on passive observations of traffic. We then apply this algorithm on packet traces collected at 55 end-hosts, and across different network environments. Our results show that 45 out of the 55 users we study experience bufferbloat at least once, 40% of these users experience bufferbloat more than once per hour. In 90% of cases, buffering more than doubles RTTs, but RTTs during bufferbloat are rarely over one second. We also show that web and interactive applications, which are particularly sensitive to delay, are the applications most often affected by bufferbloat.

6.5. Measuring and Characterising User Online Activity

Participants: Omayma Belkadi, Mauricio Santoro, Anna-Kaisa Pietilainen, Renata Teixeira

The goal of our work is to identify what people are doing online (or the online user activity) from passively collected network traffic traces. Our analysis of network traffic and application information from 12 end-hosts shows that this task is challenging because there are often many applications running on each user’s device, whereas the user is only interacting with one application at a time. Our work with two master students presents the first evaluation of the set of features computable from network traffic alone that can help distinguish user activity traffic from all other traffic flows [6], [10]. We obtain ground truth on user activities and network traffic traces in a controlled setting, and complement this dataset with traces collected by the HostView monitoring tool on the devices of 12 users over several months. We develop simple heuristics to extract user activities for the HostView dataset based on the foreground application and on keyboard/mouse activity. Then, we analyze which network traffic features allow us to distinguish between online user activity and background network traffic. Features related to traffic volumes and timings show the most significant differences.

6.6. WeBrowse: a Passive Content Curation System Based on HTTP Logs

Participants: Giuseppe Scavo, Zied Ben Houidi (Alcatel-Lucent), Renata Teixeira, Stefano Traverso (Politecnico di Torino), Marco Mellia (Politecnico di Torino)

Content curation refers to the act of assisting users to identify relevant and interesting information in the overwhelming amount of online content available today. Existing curation services rely either on experts or on crowdsourcing to promote content. This work designs, implements, and evaluates WeBrowse, the first passive crowdsourced content curation system. WeBrowse requires no active user engagement to promote content. Instead, it extracts the URLs users visit from traffic traversing an ISP network to identify popular and interesting content. A key challenge to design such a passive curation system is to process network traffic in real-time to identify the small set of URLs that are interesting to users. WeBrowse contains a set of heuristics to identify the set of URLs users visit and to select the subset that are interesting, while preserving their privacy at the same time. We prototype WeBrowse and evaluate it using traces collected at a large European ISP, and in a deployment in a large campus network. We have tested and improved WeBrowse with a small number of users from September 2014 to January 2015. The plan is to announce WeBrowse to all users of the campus network early 2015 to get feedback on their experience with the system.

Available at: http://tstat.polito.it/netcurator/

MUTANT Project-Team

6. New Results

6.1. Highlights of the Year

Acoustical Society of America Best Paper Award for [20].

International Computer Music Conference (ICMC) Best Presentation Award for [19].

MuTant TEDx Talk in October 2014 on *Human-Computer Musicianship* that attracted more than 12 thousand podcasts according to organisers.

MuTant in CNRS's 2nd edition of "Les Fondamentales" Science and Society event in Grenoble, in a session dedicated to Science and Music on the same Score.

MuTant Participation in the 2014 edition of *Futur en Seine* festival and showcased collaboration with Orchestre de Paris in a public event.
BEST PAPER AWARD :
[19] **International Computer Music Conference**. C. TRAPANI, J. ECHEVESTE.

6.2. Time-Coherency of Bayesian Priors for Sequential Alignment

In the context of Philippe Cuvillier's PhD project, we aim at increasing the robustness of machine listening in situations where observations from the external environment are extremely noisy or incoherent.

Recent results propose a novel insight to the problem of duration modeling for Information Retrieval problems where a discrete sequence of events is estimated from a time-signal using Bayesian models. Since the duration of each event is unknown, a major issue is setting the right Bayesian prior on each of them. Hidden Semi-Markov models (HSMM) allow choosing explicitly any probability distribution for the durations but learning these statistically is a non-parametric problem. In absence of huge training data sets, most algorithms rely on regularization techniques such as choosing parametric classes of distributions but the justifications of such techniques are often heuristics.

Among the numerous application domains of HMM-like paradigms, music-to-audio alignment brings two interesting properties. Firstly, a music score informs of the ordering among events. Secondly, it assigns to each event a nominal duration. For alignment tasks the Markov models conveniently model the ordering with *transient chains*. But the modeling of these nominal durations is a crucial and undermined problematic. This work investigates the relationship of this prior information of duration with the Bayesian priors of a HSMM. Theoretical insights are obtained through the study of the *prior state probability* of transient semi-Markov chains. Whereas ergodic chain and their convergence to an equilibrium probability are well studied, transient chains constitute an undermined case but of prime importance for real-time inference on HSMM.

On the first hand we prove that the non-asymptotical evolution of the state probability has some particular behaviors if the Bayesian priors fulfill several precise conditions, derived from statistical properties like the hazard rate and the tail decay. Then we say that a model is *time-coherent* if the evolution of the state probability respects the information of ordering and nominal lengths. This leads to several prescriptions on the design of HSMM Bayesian priors. On the other hand we get further prescriptions by comparing the Bayesian priors associated to different nominal lengths. This real-valued parameter comes with a natural ordering; we explain why this ordering among parameters is coherently modeled by some specific stochastic orderings among distributions that are standard in statistics.

Intermediate results have been reported in [12], [13]. This worked allowed the development of *Antescofo* version 0.6 released in November 2014.

6.3. Online Methods for Audio Segmentation and Clustering

Audio segmentation is an essential problem in many audio signal processing tasks, which tries to segment an audio signal into homogeneous chunks. Rather than separately finding change points and computing similarities between segments, we focus on joint segmentation and clustering, using the framework of hidden Markov and semi-Markov models. We introduced a new incremental EM algorithm for hidden Markov models (HMMs) and showed that it compares favorably to existing online EM algorithms for HMMs. Early experimental results on musical note segmentation and environmental sound clustering are promising and will be pursued further in 2015.

This project was done in the context of Alberto Bietti's MS project [26] under co-supervision of Arshia Cont (MuTant) and Francis Bach (SIERRA).

6.4. Model-based Testing an Interactive Music System

In the context of the Phd of Clément Poncelet, and in relation with the developments presented in Section 5.3 , we have been studying the application of model-based timed testing techniques to interactive music systems like Antescofo.

Several formal methods have been developed for automatic conformance testing of critical embedded software. The principle is to execute a real implementation under test (IUT) in a testing framework, black-box, by sending it carefully selected inputs and then observing and analyzing its outputs. In conformance model-based testing (MBT), the input and corresponding expected outputs are generated according to formal models of the IUT and the environment. The models of timed automata with inputs and outputs, and tools like the the Uppaal suite have been developed for extending such techniques to realtime systems [32], [31]. Several procedures have been designed for addressing the task described in Section 5.3 .

The case of IMS presents important originalities compared to other MBT applications to realtime systems. On the one hand, the time model supports several time units, including the wall clock time, measured in seconds, and the time of music scores, measured in number of beats relatively to a tempo. This situation raised several new problems for the generation of test suites and their execution. On the other hand, the formal specification of the IUT's behavior on a given score is produced automatically by a *score compiler*, using an intermediate representation. We rely on the realistic hypotheses that a mixed score specify completely the expected timed behavior of the IMS. Hence, our test method is fully automatic, in contrast with other approaches which generally require experts to write the specification manually. This workflow fits well in a music authoring workflow where scores in preparation are constantly evolving. We have been applying our tools to small benchmark made of characteristic scores, as well as to real mixed scores used in concerts, and some bugs in Antescofo have been identified. These results have been presented in the conference ICMC 2014 [18] and will be presented during the 30th ACM/SIGAPP Symposium On Applied Computing, track Software Verification and Testing [17].

6.5. Antescofo Temporal Pattern

An important enhancement has been made by the introduction of an expressive temporal pattern language [15] in *Antescofo*. Temporal patterns are used to define complex events that correspond to a combination of perceived events in the musical environment as well as arbitrary logical and metrical temporal conditions. The real time recognition of such event is used to trigger arbitrary actions in the style of event-condition-action rules.

The semantics of temporal pattern matching is defined to parallel the well-known notion of regular expression and Brzozowski's derivatives but extended to handle an infinite alphabet, arbitrary predicates, elapsing time and inhibitory conditions.

Temporal patterns are implemented by translation into a core subset of the Antescofo domain specific language. This compilation has proven efficient enough to avoid the extension of the real-time runtime of the language and has been validated with composers in actual pieces.

6.6. OpenMusic reactive Model

In collaboration with Jean Bresson, we have extended the evaluation model of OpenMusic to integrate reactive capabilities [10]. OpenMusic (OM) is a domain-specific visual programming language designed for computer-aided music composition based on Common Lisp. It allows composers to develop functional processes generating or transforming musical data. To extend OM towards reactive applications, we have proposed to integrate its demand-driven evaluation mechanism with reactive data-driven evaluations in a same and consistent visual programming framework. To this end, we have developped the first denotational semantics of the visual language, which gives account for its demand-driven evaluation mechanism and the incremental construction of programs. We then have extended this semantics to enable reactive computations in the functional graphs. The resulting language merges data-driven executions with the existing demand-driven mechanism. This integration allows for the propagation of changes in the programs, and the evaluation of graphically-designed functional expressions as a response to external events, a first step in bridging the gap between computer-assisted composition environments and real-time musical systems.

6.7. Representation of Rhythm and Quantization

Rhythmic data are commonly represented by tree structures (rhythms trees) in assisted music composition environments, such as OpenMusic, due to the theoretical proximity of such structures with traditional musical notation. We are studying the application in this context of techniques and tools for processing tree structure, which were originally developed for other areas such as natural language processing, automatic deduction, Web data ... We are particularly interested in two well established formalisms with solid theoretical foundations: term rewriting and tree automata.

The problem of rhythm transcription, or quantization, is to generate, from a timed sequence of notes (e.g. a file in MIDI format), a score in traditional music notation. The input events can come from an interpretation on a MIDI keyboard or be the result of a computation in OpenMusic. This problem arises immediately as insoluble unequivocally: we shall calibrate the system to fit the musical context, balancing constraints of precision, or of simplicity / readability of the generated scores. For this purpose, we are developing in collaboration with Slawek Staworko (LINKS, currently on leave at University of Edinburgh) for algorithms searching optimums in large sets of weighted trees (tree series), representing possible solutions to a problem quantification. A prototype has been developed and is under evaluation on real case studies. For the construction of appropriate tree series, we turn to semi-supervised systems, where the composer's interactions are predominant in the smooth process. These work have been presented in an invited talk in the workshop of the IFIP working group on term rewriting.

With Prof. Masahiko Sakai (Nagoya University, a specialist in term rewriting), we conduct a complementary work [14] on the representation of rhythmic notation. The goal is to define a structural theory as equations on trees rhythms. This approach can be used for example to generate, by transformation, different notations possible the same rate, with the ability to select in accordance with certain constraints.

MYCENAE Project-Team

6. New Results

6.1. Highlights of the Year

- Picture of the Conference poster of the 2014 SIAM annual meeting (July 7-11, Chicago, USA), adapted from [7]
- Invitation to organize the mini symposium "The stochastic brain" at the Stochastic Processes and Applications Conference (Jul 28-Aug1, Buenos-Aires, Argentina)
- Selection of the NeuroMathMod project in the framework of the Sorbonne Université Emergence 2014 call

6.2. Numerical and theoretical studies of slow-fast systems with complex oscillations

6.2.1. A multiple time scale coupling of piecewise linear oscillators: Application to a neuroendocrine system

Participants: Frédérique Clément, Mathieu Desroches, Soledad Fernández García, Maciej Krupa.

We have analyzed a four dimensional slow-fast piecewise linear system consisting of two coupled oscillators [32]. Each oscillator is a continuous slow-fast piecewise linear system with three zones of linearity. The coupling is one-way, that is, one subsystem evolves independently and is forcing the other subsystem. We have analyzed not only the qualitative behavior, but also quantitative aspects such as the period, frequency and amplitude of the oscillations. The system is used to reproduce all the features endowed in a former smooth model and reproduce the secretion pattern of the hypothalamic neurohormone GnRH along the ovarian cycle in different species.

6.2.2. Border collision bifurcations of stroboscopic maps in periodically driven spiking models

Participants: Frédérique Clément, Albert Granados Corsellas, Maciej Krupa.

In [21], we have considered a general nonautonomous hybrid system based on the integrate-and-fire model, widely used as simplified version of neuronal models and other types of excitable systems. Our assumptions are that the system is monotonic, possesses an attracting subthreshold equilibrium point, and is forced by means of a periodic pulsatile (square wave) function. In contrast to classical methods, in our approach we use the stroboscopic map (time-T return map) instead of the so-called firing map. It becomes a discontinuous map potentially defined in an infinite number of partitions. By applying theory for piecewise-smooth systems, we avoid relying on particular computations, and we develop a novel approach that can be easily extended to systems with other topologies (expansive dynamics) and higher dimensions. More precisely, we have rigorously studied the bifurcation structure in the two-dimensional parameter space formed by the amplitude of the pulse and the ratio between T and the duration of the pulse (duty cycle). We show that it is covered by regions of existence of periodic orbits given by period adding structures. The period adding structures completely describe not only all the possible spiking asymptotic dynamics but also the behavior of the firing rate, which is a devil's staircase as a function of the parameters.

6.2.3. Interpreting frequency responses to dose-conserved pulsatile input signals in simple cell signaling motifs

Participants: Richard Bertram, Patrick Fletcher, Joël Tabak [Florida State University], Frédérique Clément, Alexandre Vidal.

Many hormones are released in pulsatile patterns. This pattern can be modified, for instance by changing pulse frequency, to encode relevant physiological information. Often other properties of the pulse pattern will also change with frequency. How do signaling pathways of cells targeted by these hormones respond to different input patterns? We have asked if a given dose of hormone can induce different outputs from the target system, depending on how this dose is distributed in time [20]. We have used simple mathematical models of feedforward signaling motifs to understand how the properties of the target system give rise to preferences in input pulse pattern. We frame these problems in terms of frequency responses to pulsatile inputs, where the amplitude or duration of the pulses is varied along with frequency to conserve input dose. We have found that nonlinearity in the steady state input-output function of the system predicts the optimal input pattern. It does so by selecting an optimal input signal amplitude. Our results predict the behavior of common signaling motifs such as receptor binding with dimerization, and protein phosphorylation. The findings have implications for experiments aimed at studying the frequency response to pulsatile inputs, as well as for understanding how pulsatile patterns drive biological responses via feedforward signaling pathways.

6.2.4. Mixed-mode oscillations due to a singular Hopf bifurcation in a forest pest model

Participants: Morten Brøns [Technical University of Denmark], Mathieu Desroches, Maciej Krupa.

We have revisited a three-dimensional model of forest pest where MMOs play an important role [17]. In this model, young trees are distinguished from old trees, and the pest feeds on old trees. The pest grows on a fast scale, the young trees on an intermediate scale, and the old trees on a slow scale. We have established that the main organizing center for the shape and oscillatory patterns of the solutions is not a folded-node singularity, which does exist in the system, but rather a singular Hopf bifurcation. A combination of a singular Hopf bifurcation and a weak return mechanism, characterized by a very small change in one of the variables, determines the features of the mixed-mode oscillations. Period-doubling and saddle-node bifurcations lead to closed families (called isolas) of periodic solutions in a bifurcation corresponding to a singular Hopf bifurcation.

6.2.5. On the Dynamics of the adenylate energy system: homeorhesis versus homeostasis

Participants: Jesús M Cortés, Ildefonso M. de La Fuente, Iker Malaina, Luis Martínez, Edelmira Valero [University of Bilbao], Serafim Rodrigues [Plymouth University], Mathieu Desroches.

We have developed and analyzed a new model of the ATP-ADP-AMP biochemical system in order to understand some of the functional elements involved in the cellular energy status [18]. In this model based on a delayed differential system, the enzymatic rate equations and all the physiological kinetic parameters have been explicitly considered and experimentally tested in vitro. Our central hypothesis is that cells are characterized by changing energy dynamics (homeorhesis). The results have shown that the adenylate energy charge (AEC) presents stable transitions between steady states and periodic oscillations and, in agreement with experimental data these oscillations range within the narrow AEC window. Furthermore, the model shows sustained oscillations in the Gibbs free energy and in the total nucleotide pool.

6.2.6. Adaptative algorithms for the simulation of slow-fast coupled oscillators in networks

Participants: Frédérique Clément, Marie Postel, Alexandre Vidal.

The numerical simulation of a slow fast system is usually performed using an explicit scheme with an adaptive time step, in order to preserve the numerical accuracy during the fast dynamic events. In the case of large sized networks of coupled slow-fast systems, one need to use the same very small time step for all components of the network, since the integration is performed simultaneously on the whole network. We have proposed a new algorithm based on a dynamic split of the network components, in the framework of symplectic integrators [40], and applied it to a model describing the intracellular calcium oscillations in a network of embryonic GnRH neurons [9]. At each time step, the systems currently in the fast dynamic parts, are identified from their distance to the fast manifold. These components are accordingly integrated using a small time step, while a larger time step is used for the remaining of the network (cf poster abstract in the CANUM 2014 conference). Although the CPU time saving is proportional to the time constant ratio between the slow and fast dynamics, it hardly compensates the drop in the convergence order as the size of the network increases.

6.3. Non conservative transport equations for cell population dynamics

6.3.1. Adaptive mesh refinement strategy for a nonconservative transport problem

Participants: Benjamin Aymard, Frédérique Clément, Marie Postel.

In the framework of transport equations it is usual to need long time simulations, and therefore large physical domains to cover a phenomenon. On the other hand it can happen that only a small time varying portion of the domain is interesting. This motivates the use of adaptivity for the spatial discretization. Biological models involving cell development are often nonconservative to account for cell division. In that case the threshold controlling the spatial adaptivity may have to be time-dependent in order to keep up with the progression of the solution. In [16], we tackle the difficulties arising when applying a Multiresolution method to a transport equation with discontinuous fluxes modeling localized mitosis. The analysis of the numerical method is performed on a simplified model and numerical scheme. An original threshold strategy is proposed and validated thanks to extensive numerical tests. It is then applied to a biological model in both cases of distributed and localized mitosis.

6.3.2. Calibration of a multiscale model for cell dynamics

Participants: Benjamin Aymard, Frédérique Clément, Marie Postel, Kim Long Tran.

In the framework of the PhD of Benjamin Aymard and the master training of Kim Long Tran, we have tackled the issue of the numerical calibration of our multiscale model of cell populations in ovarian follicles, in collaboration with Danielle Monniaux (INRA Tours). The strategy has consisted in designing quantitative specifications from the available biological knowledge, most of which fall within the field of cell population kinetics (e.g. growth fraction, mitotic index ...), and translating them into constraints on the model parameters, as well as in performing a detailed a priori analysis of the properties of the mathematical functions entering the model equations. Using visualization approaches appropriate both for following the trajectory of a given ovarian follicle with time and comparing the follicles together, we have confronted the model outputs on different levels (from the local cell density to the overall cell number) to the corresponding specifications. We have been able to reproduce instances of the selection process occurring within a cohort of terminally growing follicles. To enable one to do systematic explorations of the model behavior in different parameter configurations associated with either physiological (e.g. species-specific ovulation number) or pathological situations (dysovulation), we have undertaken a reduction approach inspired from [41]. We have generalized these results by relaxing some simplifying assumptions to account for some important features of the original model as the distinction between different phases in the cell division cycle.

6.4. Macroscopic limits of stochastic neural networks and neural fields

6.4.1. Pulsatile localized dynamics in delayed neural-field equations in arbitrary dimension

Participants: Jonathan Touboul, Grégory Faye [EHESS].

Neural field equations are integro-differential systems describing the macroscopic activity of spatially extended pieces of cortex. In such cortical assemblies, the propagation of information and the transmission machinery induce communication delays, due to the transport of information (propagation delays) and to the synaptic machinery (constant delays). We have investigated the role of these delays on the formation of structured spatiotemporal patterns for these systems in arbitrary dimensions [19]. We have focused on localized activity, either induced by the presence of a localized stimulus (pulses) or by transitions between two levels of activity (fronts). Linear stability analysis allows to reveal the existence of Hopf bifurcation curves induced by the delays, along different modes that may be symmetric or asymmetric. We show that instabilities strongly depend on the dimension, and in particular may exhibit transversal instabilities along invariant directions. These instabilities yield pulsatile localized activity, and depending on the symmetry of the destabilized modes, either produce spatiotemporal breathing or sloshing patterns.

6.4.2. Limits and dynamics of randomly connected neuronal networks

Participants: Cristóbal Quiñinao [CIRB], Jonathan Touboul.

Networks of the brain are composed of a very large number of neurons connected through a random graph and interacting after random delays that both depend on the anatomical distance between cells. In order to comprehend the role of these random architectures on the dynamics of such networks, we have analyzed the mesoscopic and macroscopic limits of networks with random correlated connectivity weights and delays [35]. We have addressed both averaged and quenched limits, and shown propagation of chaos and convergence to a complex integral McKean-Vlasov equations with distributed delays. We have then instantiated a completely solvable model illustrating the role of such random architectures in the emerging macroscopic activity. We have particularly focused on the role of connectivity levels in the emergence of periodic solutions.

6.4.3. The propagation of chaos in neural fields

Participant: Jonathan Touboul.

We have considered the problem of the limit of bio-inspired spatially extended neuronal networks including an infinite number of neuronal types (space locations), with space-dependent propagation delays modeling neural fields [24]. The propagation of chaos property is proved in this setting under mild assumptions on the neuronal dynamics, valid for most models used in neuroscience, in a mesoscopic limit, the neural-field limit, in which we can resolve the quite fine structure of the neuron activity in space and where averaging effects occur. The mean-field equations obtained are of a new type: they take the form of well-posed infinite-dimensional delayed integro-differential equations with a nonlocal mean-field term and a singular spatio-temporal Brownian motion. We have also shown how these intricate equations can be used in practice to uncover mathematically the precise mesoscopic dynamics of the neural field in a particular model where the mean-field equations exactly reduce to deterministic nonlinear delayed integro-differential equations.

6.4.4. Spatially extended networks with singular multi-scale connectivity patterns

Participant: Jonathan Touboul.

In [24], we took care of a number of technical difficulties arising in the description of large-scale systems that are spatially extended. The organization of neurons in space (within cortical columns) and their interactions (fully connected networks) were relatively far from what is known of the anatomy of neuronal networks. In [25], we have further taken into account the fine and macroscopic structure of the cortex, which is a very large network characterized by a complex connectivity including at least two scales. On the microscopic scale, the interconnections are non-specific and very dense, while macroscopic connectivity patterns connecting different regions of the brain at larger scale are extremely sparse. This motivates to analyze the behavior of networks with multiscale coupling, in which a neuron is connected to its $v(N)$ nearest-neighbors where $v(N) = o(N)$, and in which the probability of macroscopic connection between two neurons vanishes. These are called singular multi-scale connectivity patterns. We have introduced a class of such networks and derived their continuum limit. We show convergence in law and propagation of chaos in the thermodynamic limit. The limit equation obtained is an intricate non-local McKean-Vlasov equation with delays which is universal with respect to the type of micro-circuits and macro-circuits involved.

6.4.5. Index Distribution of the Ginibre Ensemble

Participants: Romain Allez [Stastlab, Cambridge University], Gilles Wainrib [ENS], Jonathan Touboul.

Complex systems, and in particular random neural networks, are often described by randomly interacting dynamical systems with no specific symmetry. In that context, characterizing the number of relevant directions necessitates fine estimates on the Ginibre ensemble. We have computed analytically the probability distribution of the number of eigenvalues N_R with modulus greater than R (the index) of a large $N \times N$ random matrix in the real or complex Ginibre ensemble [15]. We have shown that the fraction $N_R/N = p$ has a distribution scaling as $exp(-\beta N^2 \psi_R(p))$ with $\beta = 1$ (respectively $\beta = 1/2$) for the complex (resp. real) Ginibre ensemble. For any $p \in [0, 1]$, the equilibrium spectral densities as well as the rate function $\psi_R(p)$ are explicitly derived. This function displays a third order phase transition at the critical (minimum) value $p^*_R = 1 - R^2$, associated to a phase transition of the Coulomb gas. We have deduced that, in the central regime, the fluctuations of the index N_R around its typical value $p^*_R N$ scale as $N^{1/3}$.

6.4.6. The heterogeneous gas with singular interaction: Generalized circular law and heterogeneous renormalized energy

Participants: Luis-Carlos Garcia Del Molino, Khashayar Pakdaman [Institut Jacques Monod], Jonathan Touboul.

We have introduced and analyzed d dimensional Coulomb gases with random charge distribution and general external confining potential [23]. Our long term motivation is to understand the spectrum of random matrices with non identical distributions, for instance with independent elements with distinct statistics. We have shown that these gases satisfy a large deviation principle. The analysis of the minima of the rate function (which is the leading term of the energy) reveals that at equilibrium, the particle distribution is a generalized circular law (i.e. with spherical support but non-necessarily uniform distribution). In the classical electrostatic external potential, there are infinitely many minimizers of the rate function. The most likely macroscopic configuration is a disordered distribution in which particles are uniformly distributed (for $d = 2$, the circular law), and charges are independent of the positions of the particles. General charge-dependent confining potentials unfold this degenerate situation: in contrast, the particle density is not uniform, and particles spontaneously organize according to their charge. In that picture the classical electrostatic potential appears as a transition at which order is lost. Sub-leading terms of the energy are derived: we show that these are related to an operator, generalizing the Coulomb renormalized energy, which incorporates the heterogeneous nature of the charges. This heterogeneous renormalized energy informs us about the microscopic arrangements of the particles, which are non-standard, strongly depending on the charges, and include progressive and irregular lattices.

MYRIADS Project-Team

5. New Results

5.1. Highlights of the Year

- The Contrail project coordinated by Christine Morin received the "Excellent" grade at its final review held on March 14th, 2014 in Brussels.
- Anne-Cécile Orgerie has been awarded the Young Researcher prize of the Lyon city in November 2014.
- Christine Morin has been awarded one of the 12 "Etoile de l'Europe 2014" prizes in December 2014 for the coordination of the Contrail European project.

BEST PAPERS AWARDS :

[18] **4th International Conference on Cloud Computing and Services Science**. H. FERNANDEZ, C. STRATAN, G. PIERRE.

5.2. Dependable Cloud Computing

Participants: Jiajun Cao, Stéphane Chevalier, Gene Cooperman, Teodor Crivat, Roberto-Gioacchino Cascella, Stefania Costache, Florian Dudouet, Filippo Gaudenzi, Anna Giannakou, Yvon Jégou, Ancuta Iordache, Christine Morin, Anne-Cécile Orgerie, Edouard Outin, Nikolaos Parlavantzas, Jean-Louis Pazat, Guillaume Pierre, Aboozar Rajabi, Louis Rilling, Matthieu Simonin, Arnab Sinha, Cédric Tedeschi.

5.2.1. Deployment of distributed applications in a multi-provider environment

Participants: Roberto-Gioacchino Cascella, Stefania Costache, Florian Dudouet, Filippo Gaudenzi, Yvon Jégou, Christine Morin, Arnab Sinha.

The move of users and organizations to Cloud computing will become possible when they are able to exploit their own applications, applications and services provided by cloud providers, as well as applications from third party providers in a trustful way on different cloud infrastructures. In the framework of the Contrail European project [2] [46], we have designed and implemented the Virtual Execution Platform (VEP) service in charge of managing the whole life cycle of OVF distributed applications under Service Level Agreement rules on different infrastructure providers [47]. In 2013, we designed the CIMI inspired REST-API for VEP 2.0 with support for Constrained Execution Environment (CEE), advance reservation and scheduling service, and support for SLAs [56], [55] [57]. We integrated support for delegated certificates and developed test scripts to integrate the Virtual Infrastructure Network (VIN) service. VEP 1.1 was slightly modified to integrate the usage control (Policy Enforcement Point (PEP)) solution developed by CNR. The CEE management interface was developed during 2013 and is available through the graphical API as well as through the RESTful API.

5.2.2. Checkpointing for multi-cloud environments

Participants: Jiajun Cao, Gene Cooperman, Christine Morin, Matthieu Simonin.

Most cloud platforms currently rely on each application to provide its own fault tolerance. A uniform mechanism within the cloud itself serves two purposes: (a) direct support for long-running jobs, which would otherwise require a custom fault-tolerant mechanism for each application; and (b) the administrative capability to manage an over-subscribed cloud by temporarily swapping out jobs when higher priority jobs arrive.

We propose ([31]) a novel *Checkpointing as a Service* approach, which enables application checkpointing and migration in heterogeneous cloud environments. Our approach is based on a non-invasive mechanism to add fault tolerance to an existing cloud platform *after the fact*, with little or no modification to the cloud platform itself. It achieves its cloud-agnostic property by using an external checkpointing package, independent of the target cloud platform. We implemented a prototype of the service on top of both OpenStack and Snooze IaaS clouds. We conducted a preliminary performance evaluation using the Grid'5000 experimentation platform.

5.2.3. Towards a distributed cloud inside the backbone

Participants: Anne-Cécile Orgerie, Cédric Tedeschi.

The DISCOVERY proposal currently in phase of construction and lead by Adrien Lèbre from the ASCOLA team, and currently on leave at Inria aims at designing a distributed cloud, leveraging the resources we can find in the network's backbone. [0]

In this context, and in collaboration with ASCOLA and ASAP teams, we started the design of an overlay network whose purpose is to be able, with a limited cost, to locate geographically-close nodes from any point of the network. The design, implementation, and experimentation of the overlay has been described in an article published in 2014 [22].

5.2.4. A multi-objective adaptation system for the management of a Distributed Cloud

Participants: Yvon Jégou, Edouard Outin, Jean-Louis Pazat.

In this project, we consider a "Distributed Cloud" made of multiple data/computing centers interconnected by a high speed network. A distributed Cloud is neither a usual Cloud built around a single data center, nor a Cloud Federation interconnecting different data centers owned and run by different administrative entities. Moreover, in the Cloud organization targeted here, the network capabilities can be dynamically configured in order to apply optimizations to guarantee QoS for streaming or negotiated bandwidth for example. Due to the dynamic capabilities of the Clouds, often referred to as elasticity, there is a strong need to dynamically adapt both platforms and applications to users needs and environmental constraints such as electrical power consumption.

We address the management of the Distributed Cloud in order to consider both optimizations for energy consumption and for users' QoS needs. The objectives of these optimizations will be negotiated as contracts on Service Level Agreement (SLA). A special emphasis will be put on the distributed aspect of the platform and include both servers and network adaptation capabilities. The design of the system will rely on self-* techniques and on adaptation mechanisms at any level (from IaaS to SaaS). The MAPE-k framework (Monitor-Analysis-Planning-Execution based on knowledge) will be used for the implementation of the system. The technical developments are based on the Openstack framework.

This work is done in cooperation with the DIVERSE team and in cooperation with Orange under the umbrella of the B-COM Technology Research Center.

5.2.5. Multi-cloud application deployment in ConPaaS

Participants: Stéphane Chevalier, Teodor Crivat, Guillaume Pierre.

We extended ConPaaS to support the deployment of smartphone backend applications in mobile operators' base stations. The motivation is to reduce the latency compared to a traditional deployment where the backend is located in an external cloud. This requires building a lightweight infrastructure which allows one to easily create containers that can be seamlessly migrated (roaming). A publication on this topic will appear in 2015 [23].

5.2.6. Application Performance Modeling in Heterogeneous Cloud Environments

Participants: Ancuta Iordache, Guillaume Pierre.

Heterogeneous cloud platforms offer many possibilities for applications for make fine-grained choice over the types of resources they execute on. This opens for example opportunities for fine-grained control of the tradeoff between expensive resources likely to deliver high levels of performance, and slower resources likely to cost less. We designed a methodology for automatically exploring this performance vs. cost tradeoff when an arbitrary application is submitted to the platform. Thereafter, the system can automatically select the set of resources which is likely to implement the tradeoff specified by the user. We significantly improved the speed at which the system can characterize the performance of an arbitrary application. A publication on this topic is currently under review.

[0]The DISCOVERY website: http://beyondtheclouds.github.io

5.2.7. Dynamic reconfiguration for multi-cloud applications

Participants: Nikolaos Parlavantzas, Aboozar Rajabi.

In the context of the PaaSage European project, we are working on model-based self-optimisation of multi-cloud applications. In particular, we are developing a dynamic adaptation system, capable of transforming the currently running application configuration into a target configuration in a cost-effective and safe manner. In 2014, we have defined the architecture of the adaptation system and produced a first prototype[30].

5.2.8. Self-adaptable Monitoring for Security in the Cloud

Participants: Anna Giannakou, Christine Morin, Jean-Louis Pazat, Louis Rilling.

We aim at designing a self-adaptable system for security monitoring in clouds. The considered system should cope with the dynamic nature of virtual infrastructures in clouds and have a minimal impact on performance. In 2014, we studied the state of the art in cloud security monitoring, which is composed of various approaches for intrusion detection systems (IDS), based on traditional IDS techniques such as signature-based detection and anomaly-based detection.

As a first step towards our goal of making self-adaptable a complete security monitoring architecture for cloud environments, we defined a simple initial monitoring scenario for identifying the impact of the dynamicity of a cloud architecture on the intrusion detection process. In this scenario, the security monitoring infrastructure is composed of two network IDS instances, which are used to monitor the virtual infrastructures network traffic of two cloud clients (one virtual infrastructure per client), and also eventually monitor the physical infrastructure (that is the operator's infrastructure). The virtual network traffic in each host machine is monitored by only one of the IDS instances, so that the IDS instances must be adapted to topology changes (such as migration of VMs) in the cloud environment. The adaptation process includes updates of the rules configured in the instance (deletion or creation).

In 2014, we built our testbed based on OpenStack technology for the underlying IaaS cloud platform and Snort for the network IDS. At this point the testbed consists of only five machines (on the Grid'5000 platform) but we aim to increase the number of host machines and deploy more VMs for having a more realistic representation of a production network. This will allow us to study performance issues and also more complex security monitoring setups. Our goal is also to enable monitoring of other elements, such as resource usage (both per host and per VM) on the cloud provider side.

5.2.9. Fog Computing

Participant: Jean-Louis Pazat.

The concept of "Fog Computing" is currently developed on the idea of hosting instances of services, not on centralized datacenters (i.e. the "Cloud"), but on a highly distributed infrastructure: the Internet Edge (i.e. the "Fog"). This infrastructure consists in geographically distributed computing resources with relatively small capabilities. Compared with datacenters, a "Fog" infrastructure is able to offer to Service Providers a shorter distance from the service to the user but with the same flexibility of software deployment and management.

This work focus on the problem of resource allocation in such infrastructure when considering services in the area of Internet of Things, Social Networks or Online Gaming. For such use-cases, service-to-user latency is a critical parameter for the quality of experience. Optimizing such parameter is an objective for the platform built on top of the Fog Infrastructure that will be dedicated to the deployment of the considered service. In order to achieve such a goal, the platform needs to select some strategies for the allocation of network and computing resources, based on the initial requirements for the service distribution.

We first focus on the formal expression of these requirements, by considering first the requirements provided by a Service Operator to the "Fog" Infrastructure (required computing resources, minimal quality of experience (QoE) level, etc.). The resource allocation strategies should also take into account the topology of the "Fog" Infrastructure, the heterogeneous capabilities of the equipments and of the underlying network. Based on this information, strategies and algorithms for resource allocation should be designed that will participate in the process of building an efficient platform for the service distribution. Evaluation of this efficiency will be an important process to justify the relevance of the strategies.

This work is part of Bruno Stevant's PhD thesis that began in December 2014. It is done in cooperation with the REOP team, Institut Mines telecom/IRISA.

5.3. Heterogeneous Resource Management

Participants: Eliya Buyukkaya, Djawida Dib, Eugen Feller, Christine Morin, Nikolaos Parlavantzas, Guillaume Pierre.

5.3.1. Cross-resource scheduling in heterogeneous cloud environments

Participants: Eliya Buyukkaya, Guillaume Pierre.

Allocating resources to applications in a heterogeneous cloud environment is harder than in a homogeneous environment. In a heterogeneous cloud some rare resources are more precious than others, and should be treated carefully to maximize their utilization. Similarly, applications may request groups of resources that exhibit certain inter-resource properties such as the available bandwidth between the assigned resources. We are currently investigating scheduling algorithms for handling such scenarios.

5.3.2. Maximizing private cloud provider profit in cloud bursting scenarios

Participants: Djawida Dib, Christine Morin, Nikolaos Parlavantzas.

Current PaaS offerings either provide no support for SLA guarantees or provide limited support targeting a restricted set of application types. To overcome this limitation, we have developed an open, cloud-bursting PaaS system, called Meryn, designed to be easily extensible to host new application types. The system integrates a decentralized optimization policy that maximises the PaaS provider profit, taking into account the payment of penalties incurred when quality guarantees are unsatisfied. The system was implemented and evaluated on the Grid5000 testbed using batch and MapReduce workloads. The results demonstrated the effectiveness of the policy in increasing provider profit [16] This work was part of Djawida Dib's PhD thesis [10] defended in July 2014.

5.3.3. Data life-cycle management in clouds

Participants: Eugen Feller, Christine Morin.

Infrastructure as a Service (IaaS) clouds provide a flexible environment where users can choose and control various aspects of the machines of interest. However, the flexibility of IaaS clouds presents unique challenges for storage and data management in these environments. Users use manual and/or ad-hoc methods to manage storage and data in these environments. FRIEDA is a Flexible Robust Intelligent Elastic Data Management framework that employs a range of data management strategies approaches in elastic environments. This year, our work carried out in the context of the DALHIS associate team [0], was focused on the extended design and evaluation of the FRIEDA data management system. FRIEDA was tested to work on Amazon EC2 resources. In addition, we layered a commandline utility atop FRIEDA that allows users to plug-in applications to run in FRIEDA. These tools have been adopted by the LBL-ATLAS group to run their experiments on Amazon [29].

5.4. Energy-efficient Resource Infrastructures

Participants: Maria Del Mar Callau Zori, Alexandra Carpen-Amarie, Bogdan Florin Cornea, Ismael Cuadrado Cordero, Djawida Dib, Eugen Feller, Sabbir Hasan Rochi, Yunbo Li, Christine Morin, Anne-Cécile Orgerie, Jean-Louis Pazat, Guillaume Pierre, Lavinia Samoila.

5.4.1. Energy-efficient IaaS clouds

Participants: Alexandra Carpen-Amarie, Christine Morin, Anne-Cécile Orgerie.

[0] http://project.inria.fr/dalhis

Energy consumption has always been a major concern in the design and cost of data centers. The wide adoption of virtualization and cloud computing has added another layer of complexity to enabling an energy-efficient use of computing power in large-scale settings. Among the many aspects that influence the energy consumption of a cloud system, the hardware-component level is one of the most intensively studied. However, higher-level factors such as virtual machine properties, their placement policies or application workloads may play an essential role in defining the power consumption profile of a given cloud system. In this work, we explored the energy consumption patterns of Infrastructure-as-a-Service (IaaS) cloud environments under various synthetic and real application workloads. For each scenario, we investigated the power overhead triggered by different types of virtual machines, the impact of the virtual cluster size on the energy-efficiency of the hosting infrastructure and the tradeoff between performance and energy consumption of MapReduce virtual clusters through typical cloud applications [45].

5.4.2. Energy-aware IaaS-PaaS co-design

Participants: Maria Del Mar Callau Zori, Alexandra Carpen-Amarie, Djawida Dib, Anne-Cécile Orgerie, Guillaume Pierre, Lavinia Samoila.

The wide adoption of the cloud computing paradigm plays a crucial role in the ever-increasing demand for energy-efficient data centers. Driven by this requirement, cloud providers resort to a variety of techniques to improve energy usage at each level of the cloud computing stack. However, prior studies mostly consider resource-level energy optimizations in IaaS clouds, overlooking the workload-related information locked at higher levels, such as PaaS clouds. We argue that cross-layer cooperation in clouds is a key to achieving an optimized resource management, both performance and energy-wise. To this end, we claim there is a need for a cooperation API between IaaS and PaaS clouds, enabling each layer to share specific information and to trigger correlated decisions. We identified the drawbacks raised by such co-design objectives and discuss opportunities for energy usage optimizations. A position paper has been published on these aspects [15]. Ongoing work is currently conducted in order to quantify the actual possible gains both energy and performance-wise for this IaaS-PaaS co-design approach.

5.4.3. Energy-efficient and network-aware resource allocation in Cloud infrastructures

Participants: Ismael Cuadrado Cordero, Christine Morin, Anne-Cécile Orgerie.

Cloud computing is increasingly becoming an essential component for Internet service provision, yet at the same time its energy consumption has become a key environmental and economic concern. It becomes urgent to improve the energy efficiency of such infrastructures. Our work aims at designing energy-efficient resource allocation for Cloud infrastructures. Yet, energy is not the only criterion to take into account at risk of losing users. A multi-criteria approach is required in this context to satisfy both users and Cloud providers.

The proposed resource allocation algorithms will take into account not only the computing resources but also the storage and networking resources. Indeed, the ever-growing appetite of new applications for network resources leads to an unprecedented electricity bill for network resources, and for these bandwidth-hungry applications, networks can become an significant bottleneck. This phenomenon is emphasized with the emergence of the big data paradigm. The designed algorithms would thus integrate the data locality dimension to optimize computing resource allocation while taking into account the fluctuating limits of network resources.

In 2014, several experiments were performed to understand and quantify networking energy consumption. These experiments include network protocol energy consumption in the devices, configuration energy consumption in switching/routing devices and associated energy consumption to real cloud computing applications (e.g. Google drive). These experiments have been performed over systems provided by Inria such as Grid'5000 and specific network devices (e.g. level 3 router for a private LAN). Based on this work, we developed an analytic model of networking energy consumption in a cloud computing environment. This analysis will serve as a basis for designing an energy-efficient architecture and related algorithms.

5.4.4. Simulating Energy Consumption of Wired Networks

Participants: Bogdan Florin Cornea, Anne-Cécile Orgerie.

Predicting the performance of applications, in terms of completion time and resource usage for instance, is critical to appropriately dimension resources that will be allocated to these applications. Current applications, such as web servers and Cloud services, require lots of computing and networking resources. Yet, these resource demands are highly fluctuating over time. Thus, adequately and dynamically dimension these resources is challenging and crucial to guarantee performance and cost-effectiveness. In the same manner, estimating the energy consumption of applications deployed over heterogeneous cloud resources is important in order to provision power resources and make use of renewable energies. Concerning the consumption of entire infrastructures, some studies show that computing resources represent the biggest part in Cloud's consumption, while others show that, depending on the studied scenario, the energy cost of the network infrastructure that links the user to the computing resources can be bigger than the energy cost of the servers. In this work, we aim at simulating the energy consumption of wired networks which receive little attention in the Cloud computing community even though they represent key elements of these distributed architectures. To this end, we are contributing to the well-known open-source simulator ns3 by developing an energy consumption module named ECOFEN. Through this tool, we have studied the energy consumption of data transfers in Clouds [19]. This work has been done in collaboration with the Avalon team from LIP in Lyon.

5.4.5. Resource allocation in a Cloud partially powered by renewable energy sources

Participants: Yunbo Li, Anne-Cécile Orgerie.

We propose here to design a disruptive approach to Cloud resource management which takes advantage of renewable energy availability to perform opportunistic tasks. To begin with, the considered Cloud is mono-site (i.e. all resources are in the same physical location) and performs tasks (like web hosting or MapReduce tasks) running in virtual machines. This Cloud receives a fixed amount of power from the regular electric Grid. This power allows it to run usual tasks. In addition, this Cloud is also connected to renewable energy sources (such as windmills or solar cells) and when these sources produce electricity, the Cloud can use it to run more tasks.

The proposed resource management system needs to integrate a prediction model to be able to forecast these extra-power periods of time in order to schedule more work during these periods. Batteries will be used to guarantee that enough energy is available when switching on a new server working exclusively on renewable energy. Given a reliable prediction model, it is possible to design a scheduling algorithm that aims at optimizing resource utilization and energy usage, problem known to be NP-hard. The proposed heuristics will thus schedule tasks spatially (on the appropriate servers) and temporally (over time, with tasks that can be planed in the future).

This work is done in collaboration with Ascola team from LINA in Nantes.

5.4.6. SLA driven Cloud Auto-scaling for optimizing energy footprint

Participants: Sabbir Hasan Rochi, Jean-Louis Pazat.

As a direct consequence of the increasing popularity of Internet and Cloud Computing services, data centers are amazingly growing and hence have to urgently face energy consumption issues. At the Infrastructure-as-a-Service (IaaS) layer, Cloud Computing allows to dynamically adjust the provision of physical resources according to Platform-as-a-Service (PaaS) needs while optimizing energy efficiency of the data center.

The management of elastic resources in Clouds according to fluctuating workloads in the Software-as-a-Service (SaaS) applications and different Quality-of-Service (QoS) end-user's expectations is a complex issue and cannot be done dynamically by a human intervention. We advocate the adoption of Autonomic Computing (AC) at each XaaS layer for responsiveness and autonomy in front of environment changes. At the SaaS layer, AC enables applications to react to a highly variable workload by dynamically adjusting the amount of resources in order to keep the QoS for the end users. Similarly, at the IaaS layer, AC enables the infrastructure to react to context changes by optimizing the allocation of resources and thereby reduce the costs related to energy consumption. However, problems may occur since those self-managed systems are related in some way (e.g. applications depend on services provided by a cloud infrastructure): decisions taken in isolation at given layer may interfere with other layers, leading whole system to undesired states.

We propose an approach driven by Service Level Agreements (SLAs) for Cloud auto-scaling. A SLA defines a formal contract between a service provider and a service consumer on an expected QoS level. The main idea of this thesis is to exploit the SLA requirements to (i) avoid the interferences between the Cloud autonomic managers by a cross-layer coordination of SLA contracts; (ii) fine-tune the resources needs according to SLA by proposing both dynamic resources provisioning for optimizing the energy footprint and dynamic reconfiguration at the SaaS level to optimize the expected QoS. In particular, we propose to address renewable energy in the SLA contract. The objective is twofold. First, for ecological reasons, it allows Cloud users to express their preferences about the energy provider and the nature of the energy in the data center. Then, for economic reasons, it takes advantage of renewable energy costs (expressed in the SLA) to reconfigure resource allocation and energy usage. The integration of such SLAs in each layer of the Cloud stack and their management by an autonomic manager or by the coordination of autonomic managers still remain open issues.

This work is done in collaboration with Ascola team from LINA in Nantes.

5.4.7. Simulating the impact of DVFS within SimGrid

Participants: Alexandra Carpen-Amarie, Christine Morin, Anne-Cécile Orgerie.

Simulation is a a popular approach for studying the performance of HPC applications in a variety of scenarios. However, simulators do not typically provide insights on the energy consumption of the simulated platforms. Furthermore, studying the impact of application configuration choices on energy is a difficult task, as not many platforms are equipped with the proper power measurement tools. The goal of this work is to enable energy-aware experimentations within the SimGrid simulation toolkit, by introducing a model of application energy consumption and enabling the use of Dynamic Voltage and Frequency Scaling (DVFS) techniques for the simulated platforms. We provide the methodology used to obtain accurate energy estimations, highlighting the simulator calibration phase. The proposed energy model is validated by means of a large set of experiments featuring several benchmarks and scientific applications. This work is available in the latest SimGrid release. This work is done in collaboration with the Mescal team from LIG in Grenoble.

5.5. Decentralised and Adaptive workflows

Participants: Christine Morin, Jean-Louis Pazat, Javier Rojas Balderrama, Matthieu Simonin, Cédric Tedeschi, Palakyiem Wallah.

5.5.1. Template workflows

Participants: Christine Morin, Javier Rojas Balderrama, Matthieu Simonin, Cédric Tedeschi.

In the framework of the DALHIS associate team [0], we started to combine the high-level template workflow language TIGRES [0], developed by our partner team from Lawrence Berkeley National Lab (LBL) with the workflow management system developed in the team [5]. The design of this integration and its benefits have been presented in a workhsop article [24].

5.5.2. Adaptive Workflows with Chemical Computing

Participants: Javier Rojas Balderrama, Matthieu Simonin, Cédric Tedeschi.

We are currently designing a complete programming model for the management of adaptive workflows, based on an extension of the HOCL language, in particular workflows that may evolve at run time in their shape. An article is under preparation.

5.5.3. Best-effort decentralised workflow execution

Participants: Jean-Louis Pazat, Cédric Tedeschi, Palakyiem Wallah.

[0] http://project.inria.fr/dalhis
[0] http://tigres.lbl.gov/home

We are currently proposing a simple workflow model for workflow execution in platforms with limited computing resources and services. The key idea is to devise a best-effort workflow engine that does not require a strong centralised orchestrator. Such a workflow engine relies on point-to-point cooperation between nodes supporting the execution.

5.6. Experimental Platforms

Participants: Maxence Dunnewind, Nicolas Lebreton, Julien Lefeuvre, David Margery, Eric Poupart.

5.6.1. Energy measurement

Participants: Maxence Dunnewind, Nicolas Lebreton, David Margery, Eric Poupart.

In the context of the ECO_2Clouds project, the BonFIRE infrastructure was updated. At the software layer, the complete monitoring stack was revisited so as to attribute power consumption values to all VMs running on the infrastructure and to expose this information to users. This was used by the project partners to confirm that using an eco-aware scheduler could significantly reduce eco-impact of running a distributed infrastructure.

5.6.2. BonFIRE

Participants: Maxence Dunnewind, Julien Lefeuvre, David Margery, Eric Poupart.

The project was reviewed in December 2013 during CloudCom 2013 in Bristol and rated Excellent. It has been kept in working state through our commitment to the BonFIRE foundation. The main acheivment on this topic was to evolve the cloud reservation system so as to support tracking usage using allocation blocks, as a fragment of the physical machines. Instance types can therefore have a different footprint in number of allocation blocks depending on the hardware they are scheduled on.

5.6.3. Fed4FIRE

Participants: Nicolas Lebreton, Julien Lefeuvre, David Margery.

In Fed4FIRE, two key technologies have been adopted as common protocols to enable experimenters to interact with testbeds: Slice Federation Architecture (SFA), to provision resources, and Control and Management Framework for Networking Testbeds (OMF) to control them. Here, we contributed to a proposal to secure usage of OMF and to a design to allow using BonFIRE through SFA. In 2014, the main area of work has been maintenance of the infrastructure and initial prototyping of an SFA API to BonFIRE.

NACHOS Project-Team

6. New Results

6.1. Electromagnetic wave propagation

6.1.1. Numerical study of the 1d nonlinear Maxwell equations

Participants: Loula Fézoui, Stéphane Lanteri.

The system of Maxwell equations describes the evolution of the interaction of an electromagnetic field with a propagation medium. The different properties of the medium, such as isotropy, homogeneity, linearity, among others, are introduced through *constitutive laws* linking fields and inductions. In the present study, we focus on nonlinear effects and address nonlinear Kerr materials specifically. In this model, any dielectric may become nonlinear provided the electric field in the material is strong enough. As a first setp, we consider the one-dimensional case and study the numerical solution of the nonlinear Maxwell equations thanks to DG methods. In particular, we make use of an upwind scheme and limitation techniques because they have a proven ability to capture shocks and other kinds of singularities in the fluid dynamics framework. The numerical results obtained in this preliminary study gives us confidence towards extending this work to higher spatial dimensions.

6.1.2. High order geometry conforming method for nanophotonics

Participants: Stéphane Lanteri, Claire Scheid, Jonathan Viquerat.

Usually, unstructured mesh based methods rely on tessellations composed of straight-edged elements mapped linearly from a reference element, on domains which physical boundaries are indifferently straight or curved. Such meshes represent a serious hindrance for high order finite element (FE) methods since they limit the accuracy to second order in the spatial discretization. Thus, exploiting an enhanced representation of physical geometries is in agreement with the natural procedure of high order FE methods, such as the DG method. There are several ways to account for curved geometries. One could choose to incorporate the knowledge coming from CAD in the method to design the geometry and the approximation. These methods are called *isogeometric*, and have received a lot of attention recently. This naturally implies to have access to CAD models of the geometry. On the other hand, *isoparametric* usually rely on a polynomial approximation of both the boundary and the solution. This can be added fairly easily on top of existing implementations. In the present study we focus on the latter type of method, since our goal is first to envisage the benefit of curvilinear meshes for light/matter interaction with nanoscale structures.

6.1.3. Numerical treatment of non-local dispersion for nanoplasmonics

Participants: Stéphane Lanteri, Claire Scheid, Nikolai Schmitt, Jonathan Viquerat.

When metallic nanostructures have sub-wavelength sizes and the illuminating frequencies are in the regime of metal's plasma frequency, electron interaction with the exciting fields have to be taken into account. Due to these interactions, plasmonic surface waves can be excited and cause extreme local field enhancements (surface plasmon polariton electromagnetic waves). Exploiting such field enhancements in applications of interest requires a detailed knowledge about the occurring fields which can generally not be obtained analytically. For the numerical modeling of light/matter interaction on the nanoscale, the choice of an appropriate model is a crucial point. Approaches that are adopted in a first instance are based on local (no interaction between electrons) dispersive models e.g. Drude or Drude-Lorentz. From the mathematical point of view, these models lead to an additional ordinary differential equation in time that is coupled to Maxwell's equations. When it comes to very small structures in a regime of 2 nm to 25 nm, non-local effects due to electron collisions have to be taken into account. Non-locality leads to additional, in general non-linear, partial differential equations and is significantly more difficult to treat, though. In this work, we study a DGTD method able to solve the system of Maxwell equations coupled to a linearized non-local dispersion model relevant to nanoplasmonics. While the method is presented in the general 3d case, in this preliminary stdudy, numerical results are given for 2d simulation settings.

6.1.4. Multiscale DG methods for the time-domain Maxwell equations

Participants: Stéphane Lanteri, Raphaël Léger, Diego Paredes Concha [LNCC, Petropolis, Brazil], Claire Scheid, Frédéric Valentin [LNCC, Petropolis, Brazil].

Although the DGTD method has already been successfully applied to complex electromagnetic wave propagation problems, its accuracy may seriously deteriorate on coarse meshes when the solution presents multiscale or high contrast features. In other physical contexts, such an issue has led to the concept of multiscale basis functions as a way to overcome such a drawback and allow numerical methods to be accurate on coarse meshes. The present work, which has been initiated in the context of the visit of Frédéric Valentin in the team, is concerned with the study of a particular family of multiscale methods, named Multiscale Hybrid-Mixed (MHM) methods. Initially proposed for fluid flow problems, MHM methods are a consequence of a hybridization procedure which caracterize the unknowns as a direct sum of a coarse (global) solution and the solutions to (local) problems with Neumann boundary conditions driven by the purposely introduced hybrid (dual) variable. As a result, the MHM method becomes a strategy that naturally incorporates multiple scales while providing solutions with high order accuracy for the primal and dual variables. The completely independent local problems are embedded in the upscaling procedure, and computational approximations may be naturally obtained in a parallel computing environment. In this study, a family of MHM methods is proposed for the solution of the time-domain Maxwell equations where the local problems are discretized either with a continuous FE method or a DG method (that can be viewed as a multiscale DGTD method). Preliminary results have been obtained in the 2d case for models problems.

6.1.5. HDG methods for the time-domain Maxwell equations

Participants: Alexandra Christophe-Argenvillier, Stéphane Descombes, Stéphane Lanteri.

This study is concerned with the development of accurate and efficient solution strategies for the system of 3d time-domain Maxwell equations coupled to local dispersion models (e.g. Debye, Drude or Drude-Lorentz models) in the presence of locally refined meshes. Such meshes impose a constraint on the allowable time step for explicit time integration schemes that can be very restrictive for the simulation of 3d problems. We consider here the possibility of using an unconditionally stable implicit time integration scheme combined to a HDG discretization method. As a first step, we extend our former study in [20] which was dealing with the 2d time-domain Maxwell equations for non-dispersive media.

6.1.6. HDG methods for the frequency-domain Maxwell equations

Participants: Stéphane Lanteri, Liang Li [UESTC, Chengdu, China], Ludovic Moya, Ronan Perrussel [Laplace Laboratory, Toulouse].

In the context of the ANR TECSER project, we continue our efforts towards the development of scalable high order HDG methods for the solution of the system of 3d frequency-domain Maxwell equations. We aim at fully exploiting the flexibiity of the HDG discretization framework with regards to the adaptation of the interpolation order (p-adaptivity) and the mesh (h-adaptivity). In particular, we study the formulation of HDG methods on a locally refined non-conforming tetrahedral mesh and on a non-confoming hybrid cubic/tetrahedral mesh. We also investigate the coupling between the HDG formulation and a BEM (Boundary Element Method) discretization of an integral representation of the electromagnetic field in the case of propagation problems theoretically defined in unbounded domains.

6.2. Elastodynamic wave propagation

6.2.1. Sesimic wave interaction with viscoelastic media

Participants: Nathalie Glinsky, Stéphane Lanteri, Fabien Peyrusse [Department of Mathematics, Purdue University].

This work is concerned with the development of high order DGTD methods formulated on unstructured simplicial meshes for the numerical solution of the system of time-domain elastodynamic equations. These methods share some ingredients of the DGTD methods developed by the team for the time-domain Maxwell equations among which, the use of nodal polynomial (Lagrange type) basis functions, a second order leap-frog time integration scheme and a centered scheme for the evaluation of the numerical flux at the interface between neighboring elements. A recent novel contribution is the numerical treatment of viscoelastic attenuation. For this, the velocity-stress first order hyperbolic system is completed by additional equations for the anelastic functions including the strain history of the material. These additional equations result from the rheological model of the generalized Maxwell body and permit the incorporation of realistic attenuation properties of viscoelastic material accounting for the behaviour of elastic solids and viscous fluids. In practice, we need solving 3L additional equations in 2d (and 6L in 3d), where L is the number of relaxation mechanisms of the generalized Maxwell body. This method has been implemented in 2d and 3d.

6.2.2. DG method for arbitrary heterogeneous media

Participants: Nathalie Glinsky, Diego Mercerat [CETE Méditerranée].

We have recently devised an extension of the DGTD method for elastic wave propagation in arbitrary heterogeneous media. In realistic geological media (sedimentary basins for example), one has to include strong variations in the material properties. Then, the classical hypothesis that these properties are constant within each element of the mesh can be a severe limitation of the method, since we need to discretize the medium with very fine meshes resulting in very small time steps. For these reasons, we propose an improvement of the DGTD method allowing non-constant material properties within the mesh elements. A change of variables on the stress components allows writing the elastodynamic system in a pseudo-conservative form. Then, the introduction of non-constant material properties inside an element is simply treated by the calculation, via convenient quadrature formulae, of a modified local mass matrix depending on these properties. This new extension has been validated for a smoothly varying medium or a strong jump between two media, which can be accurately approximated by the method, independently of the mesh.

6.2.3. HDG method for the frequency-domain elastodynamic equations

Participants: Hélène Barucq [MAGIQUE-3D project-team, Inria Bordeaux - Sud-Ouest], Marie Bonnasse-Gahot, Julien Diaz [MAGIQUE-3D project-team, Inria Bordeaux - Sud-Ouest], Stéphane Lanteri.

One of the most used seismic imaging methods is the full waveform inversion (FWI) method which is an iterative procedure whose algorithm is the following. Starting from an initial velocity model, (1) compute the solution of the wave equation for the N sources of the seismic acquisition campaign, (2) evaluate, for each source, a residual defined as the difference between the wavefields recorded at receivers on the top of the subsurface during the acquisition campaign and the numerical wavefields, (3) compute the solution of the wave equation using the residuals as sources, and (4) update the velocity model by cross correlation of images produced at steps (1) and (3). Steps (1)-(4) are repeated until convergence of the velocity model is achieved. We then have to solve $2N$ wave equations at each iteration. The number of sources, N, is usually large (about 1000) and the efficiency of the inverse solver is thus directly related to the efficiency of the numerical method used to solve the wave equation. Seismic imaging can be performed in the time-domain or in the frequency-domain regime. In this work which is conducted in the framework of the Depth Imaging Partnership (DIP) between Inria and TOTAL, we adopt the second setting. The main difficulty with frequency-domain inversion lies in the solution of large sparse linear systems which is a challenging task for realistic 3d elastic media, even with the progress of high performance computing. In this context, we study novel high order HDG methods formulated on unstructured meshes for the solution of the frency-domain elastodynamic equations. Instead of solving a linear system involving the degrees of freedom of all volumic cells of the mesh, the principle of a HDG formulation is to introduce a new unknown in the form of Lagrange multiplier representing the trace of the numerical solution on each face of the mesh. As a result, a HDG formulation yields a global linear system in terms of the new (surfacic) unknown while the volumic solution is recovered thanks to a local computation on each element.

6.2.4. Multiscale DG methods for the time-domain elastodynamic equations

Participants: Marie-Hélène Lallemand Tenkès, Frédéric Valentin [LNCC, Petropolis, Brazil].

In the context of the visit of Frédéric Valentin in the team, we have initiated a study aiming at the design of novel multiscale methods for the solution of the time-domain elastodynamic equations, in the spirit of MHM (Multiscale Hybrid-Mixed) methods previously proposed for fluid flow problems. Motivation in that direction naturally came when dealing with non homogeneous anisotropic elastic media as those encountered in geodynamics related applications, since multiple scales are naturally present when high contrast elasticity parameters define the propagation medium. Instead of solving the usual system expressed in terms of displacement or displacement velocity, and stress tensor variables, a hybrid mixed-form is derived in which an additional variable, the Lagrange multiplier, is sought as representing the (opposite) of the surface tension defined at each face of the elements of a given discretization mesh. We consider the velocity/stress formulation of the elastodynamic equations, and study a MHM method defined for a heterogeneous medium where each elastic material is considered as isotropic to begin with. If the source term (the applied given force on the medium) is time independent, and if we are given a arbitrarily coarse conforming mesh (triangulation in 2d, tetrahedrization in 3d), the proposed MHM method consists in first solving a series of fully decoupled (therefore parallelizable) local (element-wise) problems defining parts of the full solution variables which are directly related to the source term, followed by the solution of a global (coarse) problem, which yields the degrees of freedom of both the Lagrange multiplier dependent part of the full solution variables and the Lagrange multiplier itself. Finally, the updating of the full solution variables is obtained by adding each splitted solution variables, before going on the next time step of a leap-frog time integration scheme. Theoretical analysis and implementation of this MHM method where the local problems are discretized with a DG method, are underway.

NANO-D Project-Team

5. New Results

5.1. Variance Analysis of ARPS-Langevin dynamics

Participants: Zofia Trstanova, Gabriel Stoltz, Stephane Redon.

In order to analyze statistical convergence speed-up that can be achieved by using Adaptively Restrained Particle Simulations (ARPS) dynamics, we proposed a formula that combines the variance of the sampled process and the algorithmic speed-up:

$$S_\sigma = S_A \frac{\sigma_0^2}{\sigma_\epsilon^2} \tag{16}$$

where S_σ is the convergence speed-up, S_A is the algorithmic speed-up, σ_0^2 is the variance of the original system and σ_ϵ^2 is the variance of the ARPS-Langevin system. This led to a need of a detailed analysis of the variance of ARPS-Langevin process. We performed many numerical simulations, from the simple one-dimensional case up to more real- istic dimer-solvent models, in order to observe the behavior of the variance and the quantitative dependence on the ARPS coefficients. For the one-dimensional case we managed to compute by using Galerkin approximations the numerical approximation of the variance. We are also studying analytically by use of standard techniques the properties of the ARPS-Langevin dynamics such as the existence of an invariant measure. We are also interested in the relationship between the variance of the Langevin dynamics and the ARPS-Langevin dynamics. We showed that for small ARPS coefficients the ARPS-Langevin process can be seen as a perturbation of a standard Langevin process by a perturbation operator that depends on the ARPS coefficient ϵ.

5.2. Parallel adaptively restrained particle simulations

Participants: Krishna Kant Singh, Stephane Redon.

We have continued our work on the development of *parallel* adaptively restrained particle simulations. We have integrated the ARPS algorithm in LAMMPS (Large-scale Atomic/ Molecular Massively Parallel Simulator). LAMMPS is a computationally efficient simulator, which contains a wide range of potentials and force fields for simulating systems like solid-state materials (metals, semiconductors), soft matter (biomolecules, polymers) and coarse-grained or mesoscopic systems.

In order to verify our implementation of ARPS in LAMMPS, we have generated a trajectory of 1 ns by simulating 108 Argon particles using the ARPS algorithm and the NVE ensemble (constant Number of particles, Volume and Energy). All the particles were placed in an orthogonal box with a side length of 17.158 angstrom. We used periodic boundary conditions with 8.5 angstrom cut-off for the Lennard-Jones potential. We used a threshold $\epsilon_r = 0.0000001$ for applying restraints and a threshold $\epsilon_f = 0.005$ for releasing restraints. The system was simulated at different step sizes: using 0.5, 1, 2, 3, 4, 5, 10, 50, 60, 70, 80 and 90 femtoseconds.

Our results show that ARPS in LAMMPS preserves the total energy during simulation (Figure 4) as well as the radial distribution function (Figure 5). We are now in the process of modifying the parallel force calculation algorithms in LAMMPS to make them incremental, i.e. make their cost proportional to the number of active particles in the simulation at a given time.

5.3. Molecular Modeling

5.3.1. The CARBON method

Participants: Sergei Grudinin, Stephane Redon, Petr Popov.

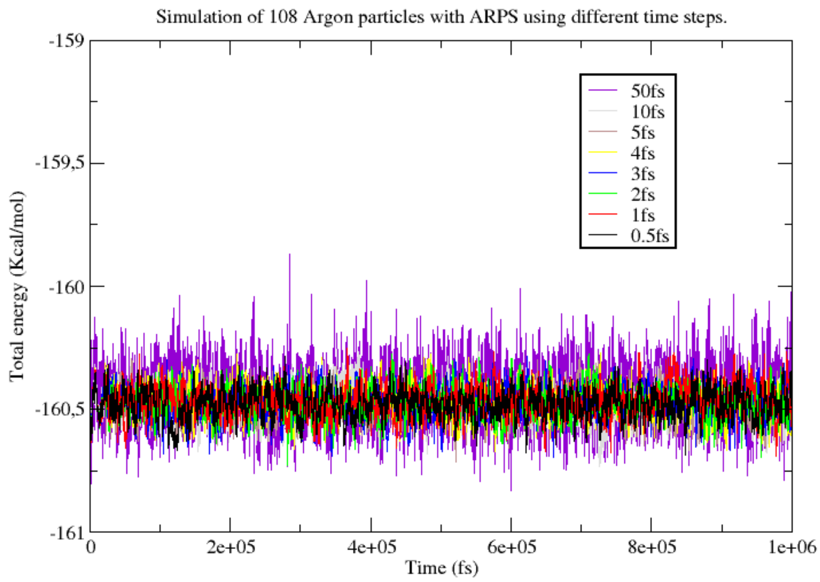


Figure 4. Energy conservation in LAMMPS using ARPS.

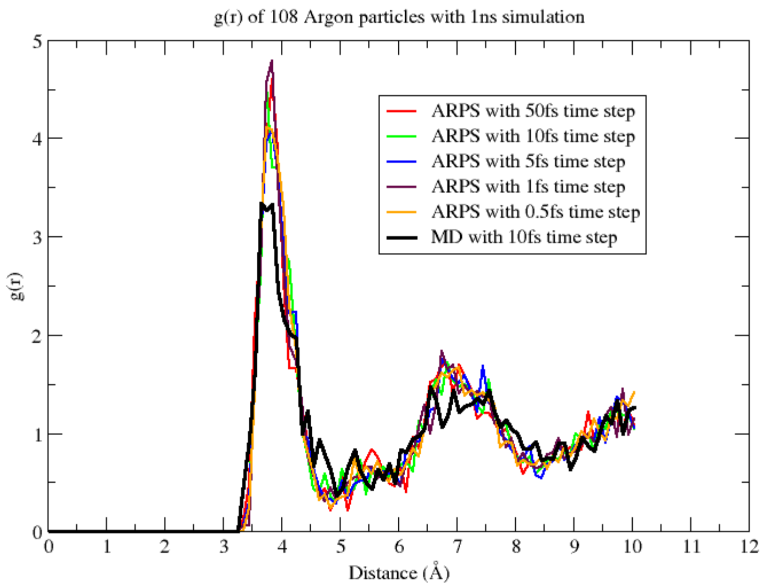


Figure 5. Preservation of the radial distribution function in LAMMPS using ARPS.

In molecular docking, various refinement algorithms are implied either to take into account flexibility of molecular complexes or to get rid of the docking artefacts, e.g. steric clashes. To address the latter problem, one possibility is to continuously minimize the energy of the complex with respect to the affine transformations, i.e. rigid transformations. Petr Popov developed a fast and efficient method called CARBON, where one considers the rigid-body optimization problem as the calculation of quasi-static trajectories of rigid bodies influenced by the inverse-inertia-weighted energy gradient. In order to determine the appropriate step-size in the direction of the net generalized force, we introduce the concept of advancement region, which is the interval of step-sizes that provide movements of the rigid body within a certain range of root mean square deviation from the initial conformation. We tested and validated CARBON on several benchmarks using both a classical force-field and a knowledge-based scoring function and demonstrated that CARBON significantly improves the quality of docking pre- dictions an also remains stable when monomers of a molecular complex significantly overlap. CARBON will be made available as a SAMSON Element for the SAMSON software platform at http://www.samson-connect.net.

5.3.2. The KSENIA method

Participants: Petr Popov, Sergei Grudinin.

Molecular docking as an integral part of the drug discovery involves the scoring stage, where one selects the best binding candidates from the set of ligand poses. The scoring stage incorporates sophisticated scoring functions based on the empirical force-fields or the information derived from known structures of protein complexes. The latter type of scoring functions belongs to the family of the knowledge-based or statistical scoring functions. Typically, for the training of a knowledge-based scoring function, modern methods require an ensemble of generated non-native decoy structures and a computation of the reference state, which is challenging. Petr Popov developed a method that does not require neither the computation of the reference state nor the ensemble of non-native complexes. Furthermore, the developed approach fully relies on the structures of protein complexes in their native configurations. More precisely, Petr trained the knowledge-based scoring function based on sets of near-native conformations. These are composed using the fluctuations along the direction of low-frequency normal modes computed at the native configurations. The obtained scoring function is capable to distinguish the native and near-native protein-protein interactions from the non-native ones. The robustness of the method was verified on several protein-protein docking benchmarks. Our methodology can be easily adapted to the recognition of other types of molecular interactions, such as protein-ligand, protein-RNA, etc. KSENIA will be made publicly available as a part of the SAMSON software platform at http://www.samson-connect.net.

5.3.3. Optimization solvers

Participants: Petr Popov, Anatoli Juditsky, Sergei Grudinin.

To derive a knowledge-based scoring function, we map non-native and near-native molecular complexes to the vectors of descriptors in a high- dimensional space. In this space, we formulate an optimization problem to construct the scoring function in such a way, that the projection of a descriptor vector onto the scoring vector corresponds to the score of a molecular complex. The formulated problem contains the regularization term and the penalty term and might vary depending on the method applied to solve the optimization problem. Different methods provide different convergence rates and cost per operation. We implemented several modern first- and second- order optimization techniques and explored which one works the best on the given data. Namely, we tested the standard gradient descent method, the conjugate gradients method, the Nesterov method, the Fista and Fista-descent methods, and the proximal gradient method.

5.3.4. Novel Docking Criterion

Participants: Petr Popov, Sergei Grudinin.

Generally, to assess the prediction capabilities of a scoring function for protein- protein interactions, one evaluates the success rate of the scoring function on widely used protein-protein benchmarks. The percentage of correctly predicted complexes is taken as the characteristic of the scoring function. However, all existing benchmarks nowadays consists on many non-native and only few near- native conformations. However, the ability of the scoring function to distinguish a particular near-native conformation from the non-native decoys does not guarantee that the scoring function is able to distinguish another near-native conformation. The same is applied if the scoring function fails on a particular molecular complex. Thus, the success rate is not a robust criterion, since it depends on the near-native and non-native conformations presented in the benchmark. We proposed the new robust method to evaluate the predictive capability of a scoring function, which does not suffer from such drawback. The method uses the probability density function of the score computed from the set of the near-native conformations and complementary empirical distribution function of the score computed from the set on non-native conformations. We tested the criterion on the previously derived scoring functions and showed that the criterion also provides an insight on some limits and restrictions of the atom-atom distant-dependent knowledge-based scoring functions.

5.4. Flexible molecular fitting

Participants: Alexandre Hoffmann, Sergei Grudinin.

We have started a PhD on flexible molecular fitting. The first part of the PhD aims at developing a new method for non-rigid molecular fitting. The problem is the following : We have two proteins $\mathcal{P}_1$ and $\mathcal{P}_2$ and we know $d_1 : \mathbb{R}^3 \to \mathbb{R}$, the electron density of $\mathcal{P}_1$ and $(Y_k)_{k=0 \cdots N_{atoms}-1}$, the average positions of the atoms of $\mathcal{P}_2$. Assuming we can generate an artificial electron density $d_2 : \mathbb{R}^3 \to \mathbb{R}$ from $(Y_k)_{k=0 \cdots N_{atoms}-1}$, our problem is to find a transformation of the atoms $T : \mathbb{R}^3 \to \mathbb{R}^3$ that minimizes the L^2 distance between d_1 and d_2.

In image processing this problem is usually solved using the optimal transport theory, but this method assumes that both densities have the same L^2 norm, which is not necessarily the case for the fitting problem. To solve this problem, one instead starts by splitting T into a rigid transformation T_{rigid} (which is a combination of translation and rotation) and a flexible transformation $T_{flexible}$. Two classes of methods have been developed to find T_{rigid} :

- the first one uses optimization techniques such as gradient descent, and
- the second one uses Fast Fourier Transform (FFT) to compute the Cross Correlation Function (CCF) of d_1 and d_2.

We have already developed several algorithms based on the FFT to find T_{rigid} and we now want to develop an efficient algorithm to find $T_{flexible}$.

The majority of algorithms first finds the best T_{rigid} and then use Normal Mode Analysis (NMA) to improve their fitting, the problem with such a method is that one can miss the optimal solution. We aim at developing a method that uses convex optimization to find the best $T_{flexible}$ for each T_{rigid} sampled on a grid, and therefore find the best T possible on a grid.

The rest of the PhD will be focused on the improvement of the modeling of the atom's motion, by using machine learning algorithms and methods that go beyond linear NMA. We hope that such an improvement can improve the quality of the fitting method.

5.5. PEPSI-Dock : Fast predictions of putative docking poses using accurate knowledge-based potentials functions to describe interaction between proteins

Participants: Emilie Neveu, Sergei Grudinin, David Ritchie, Petr Popov.

Many biological tasks involve finding proteins that can act as an inhibitor for a virus or a bacteria, fir example. Such task requires knowledge on the structure of the complex to be formed. Protein Data Bank can help but only a small fraction of its proteins are complexes [16]. Therefore, computational docking predictions, being low-cost and easy to perform, are very attractive if they describe accurately the interactions between proteins while being fast to find which conformation will be the most probable. We have been developing a fast and accurate algorithm that combines the FFT-accelerated docking methods with the precise knowledge-based potential functions describing interactions between the atoms in the proteins.

Docking methods can be described as a two ingredients recipe. First, a certain approximation for the binding free energy needed to describe the interactions between the proteins. Second, an efficient sampling algorithm is used to find the lowest-energy conformations. Commonly, as going through all the possibilities with a realistic energy function is very costly, it is approximated with a very simple energy function. Then, a much more precise energy function is typically used to re-score the most promissing predictions.Considering the numerous local minima that can be found, it is important to use the most accurate free energy from the beginning not to miss some important docking solutions. In the *Hex* code, an exhaustive search combined with a spherical polar Fourier representation enables the fast exploration of all the conformations. By now it is still the most efficient and reliable search algorithm [21]. However, only a few types of energies have been accelerated using this technic (shape complementarity and electrostatics, for example). Knowledge-based potential functions are much more precise but have been used only at the re-scoring stage of the protein docking predictions pipeline. Thus, our aim is to take advantage of the fast exhaustive search by integrating the very-detailed knowledge-based potentials into the *Hex* exhaustive search method.

We have demonstrated that we can adapt the machine learning process so that the knowledge-based potentials describing atom interactions can be translated into the polynomial basis used in *Hex*. Then, the knowledge-based scores are calculated in *Hex* using the fast polynomial expansions accelerated by the fast Fourier transform. The current evaluations of the knowledge-based scores takes more time than a shape+electrostatic representation but is still fast. More precisely, docking predictions for a single complex takes on average 5-10 minutes on a regular laptop computer. The preliminary results on the data set used for training shows significant improvements in accuracy of the method. Indeed, considering the prediction is correct if its Root Mean Square distance from the true solution is smaller than 5 Å, we currently obtain more than 50% of correct predictions rank first.

5.6. Extended Universal Force Field

Participants: Svetlana Artemova, Leonard Jaillet, Stephane Redon.

In parallel with the implementation of a Universal Force Field module in SAMSON (see Section 5.10.3), we have developed an extension of this force field to allow soft transitions for both topologies and atoms' typizations. In classical UFF topologies and atoms' typizations are set in the initialization phase and remain fixed for the entire simulation. In the proposed extension, they can vary continuously to allow the transition from one given topology to another (see Figure 6). This extended UFF combined with the interaction modeling tools already present in SAMSON allows to interactively build and modify molecules while being driven by UFF forces to ensure the relevance of the corresponding structures. The validity of this extended version of UFF was also tested on the same type of benchmarks as those used to test UFF.

5.7. Incremental Algorithms for Orbital-Free Density Functional Theory

Participants: François Rousse, Stephane Redon.

We have started a new PhD to develop incremental algorithms for electronic structure calculation.

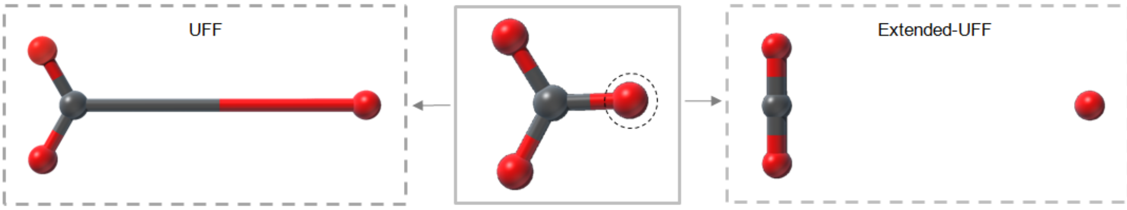


Figure 6. An oxygen atom (dashed circle) of the carbonate ion CO_3^{2-} is displaced using the interactive simulation framework in SAMSON (center). With standard UFF, the topology remains unchanged which leads to unrealistic geometries (left). With extended UFF, the covalent bond is broken forming a Carbon dioxide CO_2 and an isolated Oxygen (right).

Density Functional Theory (DFT) permits to simulate the electronic structure of a molecular system without solving the Schrödinger equation, but by finding incrementally the electronic density that minimizes the system's energy. The most used method is based on the determination of molecular orbitals. It has been shown to be an accurate method but the computation of the energy makes it too slow for the study of big systems ($> 10^3$ atoms) or dynamical ones. The Orbital-Free DFT, although less precise, is faster and can simulate the electronic density of systems up to 10^6 atoms. The aim of the PhD research is to develop new algorithms for Orbital-Free DFT that are *incremental*, i.e. whose complexity depends on the atoms that are adaptively simulated.

5.8. Robotics-inspired methods for large nanosystems

Participants: Minh Khoa Nguyen, Leonard Jaillet, Stephane Redon.

We have started a new PhD to develop robotics-inspired methods for modeling and simulating large nanosystems. Several motion planning methods issued from robotics have been successfully applied to solve problems in the field of biological molecular systems such as, including probabilistic roadmap and rapidly-exploring random trees [12]. However, large systems are still challenging due to the high number of degree of freedom. Our aim is to apply dimensionality reduction methods and/or smart conformational-space exploration techniques inspired from robotics to overcome this difficulty. The PhD topic has started since 1 Oct 2014. Reviews of the state of art and preliminary implementations have been done.

5.9. Incremental algorithms for long-range interactions

Participants: Semeho Edorh, Stephane Redon.

We have started a PhD to develop incremental algorithms for calculating long-range molecular interactions. Numerical simulation of molecular dynamics are very expensive in terms of CPU resources, especially because of the evaluation of the interaction potential. In large crystalline ionic systems, *Ewald summation* is the most popular method for computing Coulombic interactions. It rewrites the interaction potential ϕ as the sum of two terms: $\phi(r) = \phi_{dir}(r) + \phi_{rec}(r)$. The so-called "short-range" contribution ϕ_{dir} can be easily calculated in a direct space , where as the "long-range" contribution ϕ_{rec} is calculated using a Fourier transform.

Direct evaluation of the Ewald summation is an order N^2 computational problem. Over the past three decades, many techniques were developed and reduced the evaluation of the potential to an order $N \log(N)$ problem. We want to develop a new approach that can reduce the computational cost by using incremental algorithms. The key idea is to use, at each time step of the simulation, information that has been computed in previous steps.

5.10. Software development of SAMSON

5.10.1. Development of SAMSON Connect

Participants: Mohamed Yengui, Jocelyn Gate, Stephane Redon.

We have continued the development of SAMSON Connect, the web site that will contribute to the diffusion and promotion of SAMSON and SAMSON Elements (modules for SAMSON).

SAMSON Elements are adapted to different application domain and help users build new models, perform calculations, run interactive or offline simulations, visualize and interpret results, etc. The goal of SAMSON Connect is to bring together a set of users and developers of SAMSON Elements in all areas of nanoscience (physics, biology, chemistry, electronics, etc...). It offers a set of features available depending on the user role:

- Developers (who have obtained the SAMSON-SDK) can develop SAMSON Elements and upload them to SAMSON Connect through the tools provided.
- Users (who have obtained the SAMSON Core application) can add SAMSON Elements to their instance of SAMSON Core in one click. The download process is performed during startup of SAMSON and without outside intervention.

All users can give feedbacks, review and rate SAMSON Elements after adding them to their SAMSON Core (Figure 7).

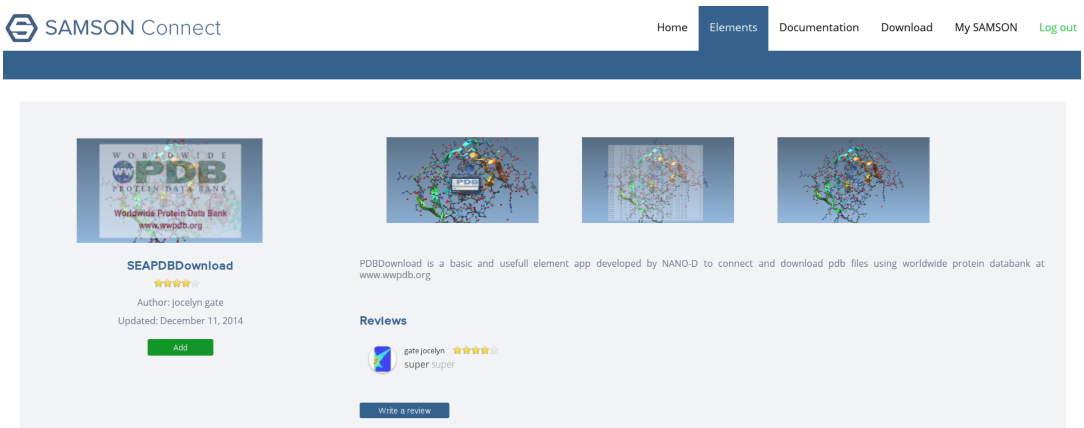


Figure 7. Screenshot of a SAMSON Element on SAMSON Connect.

SAMSON Connect also features some documentation to develop new elements for SAMSON (Figure 8).

SAMSON Connect will be available at http://samson-connect.net.

5.10.2. Deployment of SAMSON and the SAMSON SDK

Participants: Jocelyn Gate, Mohamed Yengui, Stephane Redon.

The SAMSON installer has been split in two parts: SAMSON-setup (installation of the SAMSON application, Figure 9) and SAMSON-Developer-setup (installation of the SAMSON SDK). internet. It is very useful to increase security.

Several helper tools related to SAMSON Elements management were developed to facilitate Element deployment. For example, the element packager is a tool useful for developers who want to distribute a new SAMSON Element on the SAMSON Connect platform. With this packager we can control many things: check whether the file is valid, if the SAMSON Element is readable with SAMSON, add a description file that contains useful information (name, author ID, checksum, element version, SDK version, operating system, etc.).

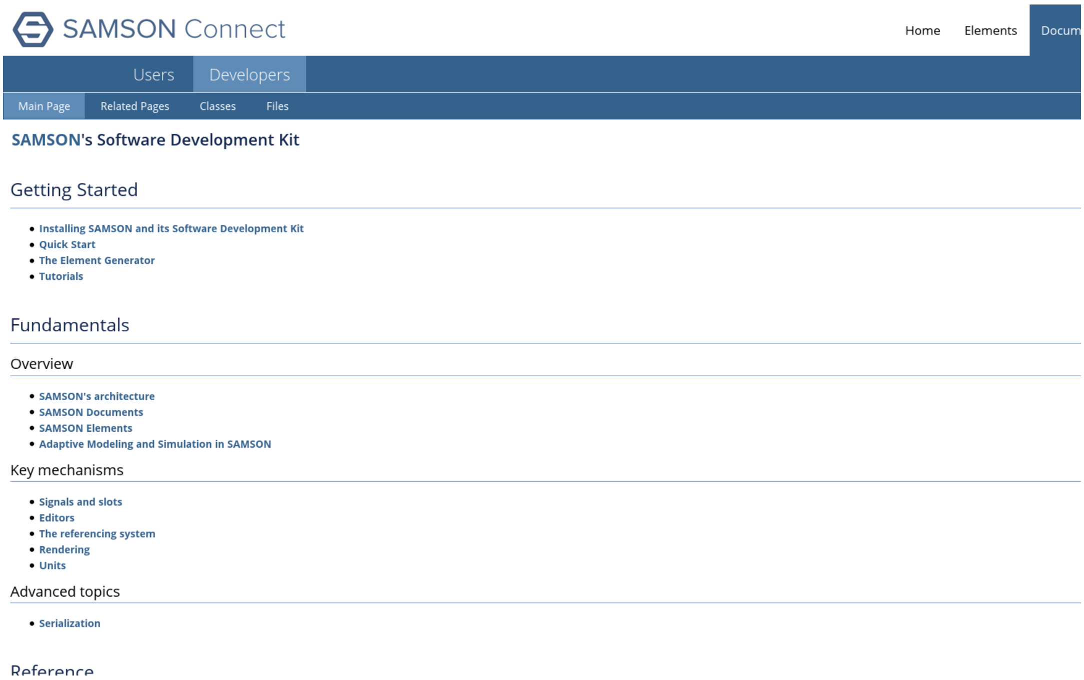


Figure 8. Screenshot of documentation on SAMSON-Connect.

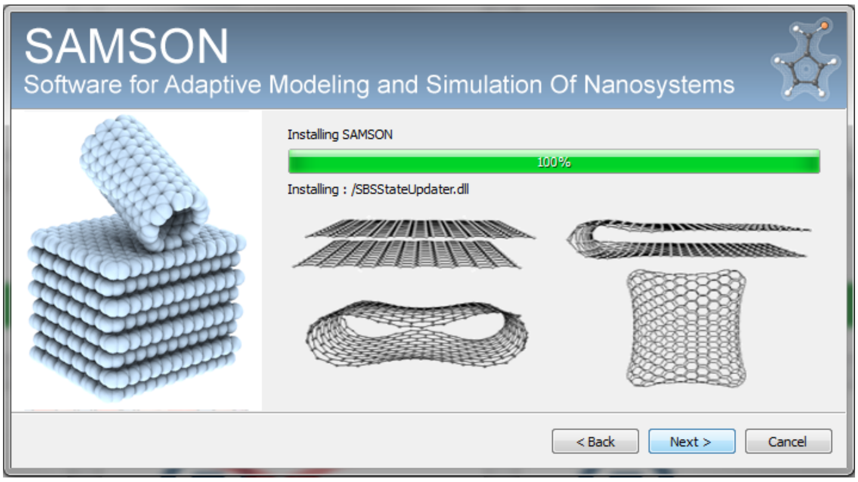


Figure 9. The SAMSON Installer

We added a service requester to SAMSON to communicate with SAMSON Connect and

- Check users/developers status
- Easily download new SAMSON Elements
- Be notified about updates

5.10.3. Universal Force Field

Participants: Svetlana Artemova, Leonard Jaillet, Stephane Redon.

We have implemented a version of the Universal Force Field (UFF) [19] in SAMSON, as a SAMSON Element embedding an interaction model. UFF is a classical force field, which can take as input almost every atom of the periodic table. Such flexibility allows to potentially use UFF on a large spectrum of systems and since its introduction, it has been applied to simulate problems involving main group compounds, organic molecules, metal complexes and has even been recently extended to MOF (Metal Organic Framework) [11]. The general energy expression for UFF as described in [19] is:

$$E_{UFF} = E_R + E_\theta + E_\phi + E_\omega + E_{vdw} + E_{el},$$

where E_R stands for bond stretching, E_θ describes angle bending, E_ϕ is dihedral angle torsion term, E_ω represents inversion, E_{vdw} stands for van der Waals interactions and E_{el} represents electrostatics (this last term is rarely considered for UFF, we do not study it neither). Forces involved in the atoms interactions can then be derived from the previous expression. Each energetic term in UFF can be computed based on simple rules deduced from a set of parameters. This set is based on the atoms' elements, their hybridization, and the overall connectivity of the molecular system.

In our implementation, we took into account several corrections and refinements that have been lately proposed in the literature for Universal Force Field. Our contribution also concerns the development of algorithms to automatically perceive the system's topology (covalent bonds and bond orders assignments). Moreover, we have introduced a method to automatically find the correct typization of the atoms. Precisely, atoms' hybridizations and oxidation states are computed, and resonance groups within or out of cycles are detected and treated. The implementation provided is computationally efficient enough to allow interactive simulation in SAMSON. The validity of the force field was tested on several groups of molecules proposed as benchmarks in the literature.

5.10.4. Integration of existing tools

Participants: Nadhir Abdellatif, Svetlana Artemova, Stephane Redon.

We have obtained funding from the Nanosciences Foundation in Grenoble to integrate in SAMSON some tools developed and used by the Grenoble community, in the form of SAMSON Elements, i.e. modules that integrate into SAMSON and may interact with SAMSON's main data graph. In particular, we have been meeting with some biologists and physicists to determine which tools and methods used (or developed) in Grenoble would be most appropriate for integration.

We integrated our first Element which is Babel, a chemical toolbox designed to "speak the many languages of chemical data", i.e. read, write and convert data files (over 110 chemical file formats) from molecular modeling, chemistry, solid-state materials, biochemistry, or related areas (see http://openbabel.org). The corresponding SAMSON element is an *app* that delegates all calculations to the Babel external executable. The app also makes it possible to import the data files to SAMSON to visualize the molecular data and proceed with other SAMSON elements.

We have also integrated Clustal, a tool for multiple sequence alignment. Thanks to Clustal's license, all source code is wrapped into the SAMSON Element (whose source code will be made available as well), and SAMSON users do not need to install Clustal separately.

5.10.5. Various

Participants: Stephane Redon, Svetlana Artemova, Marc Aubert.

- Units management was added to SAMSON. The mechanism relies on C++ template meta-programming techniques to perform dimensional analysis and automatic conversions at compile time, and has no runtime overhead. This was a significant undertaking, but one that will be very helpful to integrate in SAMSON different domains of nanoscience that have come to use different units for identical dimensions (e.g. kilocalories per mole in biology, electron volts in chemistry, etc.).
- SAMSON's reflection mechanism was improved to perform type registration and casting, and facilitate scripting and pipelining of SAMSON Elements.
- SAMSON now handles multiple documents.
- SAMSON has its own file format, which allows it to save the data graph information.
- More data graph nodes are now visible in SAMSON's data graph view.
- The split between classical and quantum interaction models was abandoned, for simplicity.
- SAMSON now handles multiple cameras.
- Selection methods have been improved, and selection is now undoable. Selections may be saved, retrieved, have boolean operations performed onto them, etc.
- The documentation of the SAMSON SDK has been improved.
- Controllers, a new type of data graph nodes, were added to SAMSON. Controllers are used to act on other data graph nodes (e.g. translate and rotate models).
- The object lifecycle of SAMSON was improved.
- SAMSON now has a mechanism for serialization.
- SAMSON now has preferences (e.g. for rendering).
- Existing parsers for input and output of molecular information in SAMSON have been improved and accelerated, and property windows for these parsers have been added.
- The Lennard-Jones potential has been added as an interaction model to SAMSON.
- A new editor for adding atoms corresponding to a chemical formula (in disorder) has been created.
- The work on a new editor containing functional groups and frequently-used molecular patterns has been started.
- Periodic Boundary Conditions (an important concept in molecular simulations) were implemented in SAMSON.
- General code debugging and improvement has been performed
- We decided to use the Qt5 framework for shaders management, for some maintenance reasons especially. This structure implied some other type changes to adapt to Qt5, such as the vertex buffers.
- We changed the way viewports display text. It is now possible to run SAMSON on every platform (Windows, Linux and Mac) and display text, and it provides Elements programmers a simple way to add text where they want in the 3D view.

NECS Project-Team

6. New Results

6.1. Highlights of the Year

- C. Canudas de Wit serves as General Chair for the Europeen Control Conference (ECC'14), Strasbourg, France, Jul. 2014 (http://www.ecc14.eu/).
- Launch of the SPEEDD EU FP7 project in Feb. 2014.
- Launch of the COMFORT project, which supports the associate Team between Inria project-team NeCS and the Berkeley University project PATH (http://necs.inrialpes.fr/pages/projects/comfort.php).
- Launch of the LOCATE-ME Persyval project (Apr. 2014 to Aug. 2015) in collaboration with the Tyrex team.
- The team has organized the Hycon2 Show day in May 2014 (http://www.inria.fr/en/centre/grenoble/calendar/hycon2-show-day-traffic-modeling-estimation-and-control).

6.2. Networked systems and graph analysis

6.2.1. Distributed solution to the network reconstruction problem

Participants: A. Kibangou [Contact person], F. Morbidi.

It has been recently shown in [45] that by collecting noise-contaminated time series generated by a coupled-oscillator system at each node of a network, it is possible to robustly reconstruct its topology, i.e. determine the graph Laplacian. Restricting ourselves to linear consensus dynamics over undirected communication networks. In [18], we have introduced a new dynamic average consensus least-squares algorithm to locally estimate these time series at each node, thus making the reconstruction process fully distributed and more easily applicable in the real world. We have also proposed a novel efficient method for separating the off-diagonal entries of the reconstructed Laplacian, and examined several concepts related to the trace of the dynamic correlation matrix of the coupled single integrators, which is a distinctive element of our network reconstruction method.

6.2.2. Distributed estimation of Laplacian eigenvalues and network robustness assessment

Participants: A. Kibangou [Contact person], T.-M. D. Tran, J. Hendrickx [Univ. Louvain-la-neuve].

As recently shown, Laplacian eigenvalues can be estimated by solving the factorization of the average consensus Matrix [46]. The problem was viewed as a constrained consensus optimization one. The main assumption was about the knowledge of the final consensus value. Indeed, estimation of the Laplacian eigenvalues can be carried out using measurements of the transient of the consensus protocol and the steady state (consensus value). In [34], we relaxed the assumptions by considering that the consensus value is only approximately known. We formulated a convex optimization, which allowed us to make use of recent well-known techniques and results dealing with convex optimization problem proposed in the literature (the Alternating Direction of Multipliers Method, ADMM), [40], [42]. Recently, we assumed that the consensus value is completely unknown and has to be found simultaneously with Laplacian eigenvalues. In such a case the problem becomes a convex combination problem where the cost function comprises two terms, one that is average consensus problem, and the rest is the consensus problem to estimate the Laplacian eigenvalues. The simulations indicate that the proposed algorithm is efficient enough to provide the nonzero distinct Laplacian eigenvalues with high accuracy. These eigenvalues are then used to assess the robustness of the graph by means of some spectral metrics, the number of spanning trees and the Kirchoff index precisely.

6.2.3. Observability and privacy preserving features in consensus networks

Participants: A. Kibangou [Contact person], C. Commault [Grenoble INP].

In [16], we have studied of observability in consensus networks modeled with strongly regular graphs or distance regular graphs. The first result consists in a Kalman-like simple algebraic criterion for observability in distance regular graphs. This criterion consists in evaluating the rank of a matrix built with the components of the Bose-Mesner algebra associated with the considered graph. Then, we have defined some bipartite graphs that capture the observability properties of the graph to be studied. In particular, we showed that necessary and sufficient observability conditions are given by the nullity of the so-called local bipartite observability graph (resp. local unfolded bipartite observability graph) for strongly regular graphs (resp. distance regular graphs). When the nullity cannot be derived directly from the structure of these bipartite graphs, the rank of the associated bi-adjacency matrix allows evaluating observability. Eventually, as a by-product of the main results we have shown that non-observability can be stated just by comparing the valency of the graph to be studied with a bound computed from the number of vertices of the graph and its diameter. Similarly non-observability can also be stated by evaluating the size of the maximum matching in the above mentioned bipartite graphs. Non-observability is strongly linked to privacy preserving feature of a given network. Indeed, when a node is neighborhood non-observable, it means that the data of the other nodes (excluding those of its neighborhood) cannot be retrieve from such a node. Therefore security efforts in order to preserve privacy of the entire network must be focused on nodes that are neighborhood-observable.

6.2.4. Average and parametric consensus

Participants: A. Kibangou [Contact person], F. Morbidi.

We have studied average consensus in wireless sensor networks with aim of providing a way to reach consensus in a finite number of steps [17]. In particular, we investigate the design of consensus protocols when, for security reasons for instance, the underlying graph is constrained to be strongly regular or distance regular. The proposed design method is based on parameters of the intersection array characterizing the underlying graph. With this protocol, at execution time, average consensus is achieved in a number of steps equal to the diameter of the graph, i.e. the smallest possible number of steps to achieve consensus. We have extended the parametric consensus protocol recently introduced by F. Morbidi, to more realistic agents modeled as double integrators and interacting over an undirected communication network. The stability properties of the new protocol in terms of the real parameter "s" are studied for some relevant graph topologies, and the connection with the notion of bipartite consensus is highlighted. The theory is illustrated with the help of two worked examples, dealing with the coordination of a team of quadrotor UAVs and with cooperative temperature measurement in an indoor environment [32].

6.3. Collaborative and distributed algorithms

6.3.1. Distributed computation methods for large-scale multidimensional data

Participants: A. Kibangou [Contact person], T.-M. D. Tran, A. de Almeida [UFC Brazil].

From Internet to large research infrastructures, the volume of data generated by our societies is continuously increasing. A deluge faced by the producers of these data as well as their users. The big data issue is a significant scientific challenge that requires deep investigations in both engineering and fundamental science. Low-rank matrix factorization has received a particular attention in recent years, since it is fundamental to a variety of mining tasks that are increasingly being applied to massive datasets. In large applications, matrix factorizations can involve matrices with billions of entries. At this massive scale, distributed algorithms for matrix factorization are essential to achieve reasonable performance [43]. However, in many disciplines, data inherently has more than two axes of variation and can be arranged as tensors (i.e. multi-way arrays). Computing tensor decompositions of multi-way datasets is particularly useful to extract hidden patterns and structure in data analytics problems. Specifically, CPD (Canonical Polyadic Decomposition) also known as PARAFAC (Parallel factor analysis) is an extension of a low rank matrix decomposition to tensors. In [26], we have introduced a fully distributed method to compute the CPD of a large-scale data tensor across a network of machines with limited computation resources. The proposed approach is based on collaboration between the machines in the network across the three modes of the data tensor. Such a multi-modal collaboration allows an essentially unique reconstruction of the factor matrices in an efficient way. We provide an analysis of the computation and communication cost of the proposed scheme and address the problem of minimizing communication costs while maximizing the use of available computation resources.

6.3.2. Collaborative source seeking

Participants: C. Canudas de Wit [Contact person], R. Fabbiano, F. Garin, Y. Gaudfrin, J. Dumon.

The problem of source localization consists in finding, with one or several agents possibly cooperating with each other, the point or the spatial region from which a quantity of interest is being emitted. Source-seeking agents can be fixed sensors, that collect and exchange some information about the signal field and try to identify the position of the source (or the smallest region in which it is included), or moving devices equipped with one or more sensors, that physically reach the source in an individual or cooperative way. This research area is attracting a rapidly increasing interest, in particular in applications where the agents have limited or no position information and GPS navigation is not available, as in underwater navigation or in cave exploration: for instance, source localization is relevant to many applications of vapor emitting sources such as explosive detection, drug detection, sensing leakage or hazardous chemicals, pollution sensing and environmental studies. Other fields of interest are sound source localization, heat source localization and vent sources in underwater field. Techniques present in literature either are based on a specific knowledge of the solution of the diffusion process, or make use of an extremum-seeking approach, exciting the system with a periodic signal so as to explore the field and collect enough information to reconstruct the gradient of the quantity of interest. Our approach [13] lies in the computation of derivatives (potentially of any order) from Poisson integrals that, for isotropic diffusive source in steady-state, whose solution satisfies the Laplace equation, allows for a gradient search with a small computation load (derivatives are computed by integrals) and without requiring any knowledge of the closed-form solution, avoiding in the same time extremum-seeking oscillations; this has the additional advantage of an intrinsic high-frequency filtering, that makes the method robust to measurement noise. We also propose a distributed version of this algorithm [28], where agents communicate in order to reconstruct gradient information from local pointwise measurements, and a control law combines the two objectives of formation control (to have a circular formation, so that measurements are taken around circle) and gradient ascent (so as to move towards the source); differently from previous literature [41], the moving agents do not need to know their absolute position, but only relative bearing angle of their neighbours.

6.4. Sensor networks: estimation and data fusion

6.4.1. Data fusion approaches for motion capture by inertial and magnetic sensors

Participants: H. Fourati [Contact person], A. Makni, A. Kibangou.

The problem of rigid body attitude estimation under external acceleration from a small inertial/magnetic sensor module containing a triaxial gyroscope, accelerometer, and magnetometer is considered [15]. We are focused on two main challenges. The first one concerns the attitude estimation during dynamic conditions, in which external acceleration occurs [30]. Although external acceleration is one of the main source of loss of performance in attitude estimation methods, this problem has not been sufficiently addressed in the literature. A quaternion based adaptive Kalman filter (q-AKF) compensating external acceleration from the residual in the accelerometer is designed. At each step, the covariance matrix associated with the external acceleration is estimated to adaptively tune the filter gain. The second challenge deals with the energy consumption issue of gyroscope for a long-term battery life of Inertial Measurement Units (IMUs). We study the way to reduce the gyro measurement acquisition by switching on/off the sensor while maintaining acceptable attitude estimation. A smart detection approach is proposed to decide whether the body is in dynamic or static motion. The efficiency of the q-AKF is investigated through numerical simulations and experimental tests, under external acceleration and parsimonious use of gyroscope. This work is described in a submitted in IEEE/ASME Transactions on Mechatronics.

6.4.2. Pedestrian dead-reckoning navigation

Participant: H. Fourati [Contact person].

We propose a foot-mounted Zero Velocity Update (ZVU) aided Inertial Measurement Unit (IMU) filtering algorithm for pedestrian tracking in indoor environment. The algorithm outputs are the foot kinematic parameters, which include foot orientation, position, velocity, acceleration, and gait phase. The foot motion filtering algorithm incorporates methods for orientation estimation, gait detection, and position estimation. A novel Complementary Filter (CF) is introduced to better pre-process the sensor data from a foot-mounted IMU containing tri-axial angular rate sensors, accelerometers, and magnetometers and to estimate the foot orientation without resorting to GPS data. A gait detection is accomplished using a simple states detector that transitions between states based on acceleration measurements. Once foot orientation is computed, position estimates are obtained by using integrating acceleration and velocity data, which has been corrected at step stance phase for drift using an implemented ZVU algorithm, leading to a position accuracy improvement. We illustrate our findings experimentally by using of a commercial IMU during regular human walking trial in a typical public building. Experiment results show that the positioning approach achieves approximately a position accuracy less than 1 m and improves the performance regarding a previous work of literature [14].

6.4.3. Sensor placement of unreliable sensors

Participants: F. Garin [Contact person], P. Frasca [U. Twente], B. Gerencsér [U. Catholique de Louvain], J. Hendrickx [U. Louvain-la-neuve].

We consider problems in which sensors have to be deployed in a given environment in such a way to provide good coverage of it. It is clear that sensor failures may deteriorate the performance of the resulting sensor network. Then, it is also natural to ask if taking into account such uncertainties changes the coverage optimization problem and leads to a different optimal solution. For simplicity, we start considering a one-dimensional problem, where sensors are to be placed on a line in such a way to optimize the disk-coverage cost. The optimal solution for reliable sensors is simply an equally-spaced configuration of the sensors. If we allow that the sensors may fail to take or communicate their measurements, this solution may instead not be optimal. In our work, we assume that sensor can fail, independently and with a same failure probability, and we aim to minimize, in expectation, the largest distance between a point in the environment and an active sensor. Our first result states that the problem at hand is equivalent to a linear program, albeit with a number of variables growing exponentially with the number of sensors. This fact allows for a computational solution that is tractable if the number of sensors is not large. Secondly, we show that for large number of sensors n, the cost of the equispaced placement is asymptotically optimal, i.e., the ratio between its cost and the optimal cost tends to 1 when n grows. By contrast, we show in that a random sensor placement has an expected cost which is larger. This work has been presented at MTNS conference [35] and is described in a submitted journal paper (see http://arxiv.org/abs/1404.7711).

6.5. Control design and co-design

6.5.1. Energy-aware networked control

Participants: C. Canudas de Wit [Contact person], F. Garin, N. Cardoso de Castro, D. Quevedo [U. Newcastle].

We have considered an event-based approach to energy-efficient management of the radio chip in the sensor node of a wireless networked control system. Indeed the radio is the main energy consumer, and intermittent data transmission allows one to reduce the use of the radio. While the existing literature in the control community on event-based control only addresses policies using two radio modes (transmitting/sleep), our work follows some considerations on the radio chip modes well-known in the communication networks literature, and introduces various radio-modes: different ‘idle’ non-transmitting modes, where only part of the radio chip is switched off (thus consuming more energy than ‘sleep’, but allowing for faster transition to transmission), and various transmitting modes, with different power levels. We propose an event-based radio-mode switching policy, which allows to perform a trade-off between energy saving and performance of the control application; to this end, a switched model describes the system, taking into account control and communication. The optimal switching policy is computed using dynamic programming, considering a cost either over an infinite time-horizon (see [36]) or over a finite receding horizon (joint work with D. Quevedo, Univ. Newcastle, Australia, described in a paper in preparation).

6.5.2. Adaptive control strategy based reference model for spacecraft motion trajectory

Participants: H. Fourati [Contact person], Z. Samigulina [Kazakh National Technical University], O. Shiryayeva [Institute of Informatics and Control Problems].

In aerospace field, the economic realization of a spacecraft is one of the main objectives which should be accomplished by conceiving the optimal propulsion system and the best control algorithms. Our work focuses on the development of a viable Adaptive Control Approach (ACA) for Spacecraft Motion Trajectory (SMT), see [19]. The proposed strategy involves the nonlinear mathematical model of SMT expressed in the central field, which is linearized by the Taylor expansion, and the second Lyapunov method to offer a high rate and unfailing performance in the functioning. The adaptive control system is composed of the cascade of adaptation loop and feedback control loop. When the spacecraft deviates from its reference trajectory model, the ACA acts on the control system to correct this deviation and follow the optimal reference trajectory. Therefore, when the states of the adjustable model are different from its reference values, then the error signal is provided as an input to the adaptation law, which contains the adaptation algorithm. The output will be the state variable feedback control matrix which will be used to calculate the new control law vector. The efficiencies of the linearization procedure and the control approach are theoretically investigated through some realistic simulations and tests under MATLAB. The steady state errors of control between the reference model and the adjustable model of SMT converge to zero. This work is described in [20].

6.5.3. Control design for hydro-electric power-plants

Participants: C. Canudas de Wit [Contact person], S. Gerwig, F. Garin, B. Sari [Alstom].

We have initiated a collaboration with Alstom on collaborative and resilient control of hydro-electric power-plants, with the CIFRE PhD thesis of Simon Gerwig. The first goal of this research is to improve performance of a hydro-electric power-plant outside its design operation conditions, by adaptive cancellation of oscillations that occur in such an operation range. Indeed, current operation of power-plants often requires to operate on a variety of conditions, often different from the ones initially considered when designing the plant. At off-design operation pressure, the hydraulic turbine exhibits a vortex rope below the runner. This vortex generates pressure fluctuations after the turbine and can excite the hydraulic pipes. Indeed the water is compressible and the pipe walls elastic, so the system can oscillate. The goal is to damp these pressure oscillations as they create vibrations in the system and can lead to damages. Our first contribution has been to model the effect of the vortex rope on the hydraulic system as an external perturbation source acting on pipes. The pipes themselves are described with equations taking into account water compressibility and pipe-wall elasticity. The resulting model is nonlinear with hyperbolic functions in the equations (analogous to high-frequency transmission lines), from which we obtain a suitably linearized model.

6.5.4. Controller for switched linear systems

Participants: H. Fourati [Contact person], Djamel. E. C. Belkhiat [U. Setif], D. Jabri [U. Setif].

We designed a robust output feedback tracking controller for a class of Switched Linear Systems (SLS) subject to external disturbances [23]. The proposed synthesis approach, based on a descriptor redundancy formulation, allows to avoid of the crossing terms appearance between the switched Proportional-Derivative (PD) controller's and the switched system's matrices. Using the multiple Lyapunov functional methods, a robust output feedback tracking performance has been formulated in terms of set of Linear Matrix Inequality (LMI). The effectiveness of the proposed synthesis procedure has been illustrated by a numerical example [24].

6.6. Transportation networks and vehicular systems

6.6.1. Traffic estimation and prediction

Participants: C. Canudas de Wit [Contact person], A. Kibangou, L. Leon Ojeda, F. Morbidi.

In the PhD thesis of Leon Ojeda, we have been concerned with the design of a methodology for the real-time multi-step ahead travel time forecasting using flow and speed measurements from an instrumented freeway. Two main methodologies have been considered. The first one, a signal-based, uses only speed measurements collected from the freeway, where a mean speed is assumed between two consecutive collection points. The travel time is forecasted using a noise Adaptive Kalman Filter (AKF) approach. The process noise statistics are computed using an online unbiased estimator, while the observations and their noise statistics are computed using the clustered historical traffic data. Forecasting problems are reformulated as filtering ones through the use of pseudo-observations built from historical data. The second one, a model-based, uses mainly traffic flow measurements. Its main appealing is the use of a mathematical model in order to reconstruct the internal state (density) in small road portions, and consequently exploits the relation between density and speed to forecast the travel time. The methodology uses only boundary conditions as inputs to a switched Luenberger state observer, based on the "Cell Transmission Model" (CTM), to estimate the road initial states. The boundary conditions are then forecasted using the AKF developed above. Consequently, the CTM model is run using the initial conditions and the forecasted boundaries in order to obtain the future evolution of densities, speeds, and finally travel time. The added innovation in this approach is the space discretization achieved: indeed, portions of the road, called "cells", can be chosen as small as desired and thus allow obtaining a finer tracking of speed variations. The developed methodologies were assessed using the city-lab GTL [31]. Features and activities of this platform are described in [39].

6.6.2. Traffic control

Participants: C. Canudas de Wit [Contact person], D. Pisarski.

The work was mainly focused on the final design of a distributed controller and its implementation to the model of the south ring of Grenoble in the context of the project Hycon2. For the sake of the controller design, a distributed optimal control method for balancing of freeway traffic density was studied. The optimization was performed in a distributed manner by utilizing the controllability properties of the freeway network represented by the Cell Transmission Model. By using these properties, the subsystems to be controlled by local ramp meters were identified. The optimization problem was then formulated as a non-cooperative Nash game. The game was solved by decomposing it into a set of two-players hierarchical and competitive games. The process of optimization employed the communication channels matching the switching structure of system interconnectivity. By defining the internal model for the boundary flows, local optimal control problems were efficiently solved by utilizing the method of Linear Quadratic Regulator. The developed control strategy was tested via numerical simulations on the macroscopic model in two scenarios for uniformly congested and transient traffic. The controller was also validated through a microscopic simulations with the use of Aimsun software. The controller was implemented through Matlab under which a relevant program simulating distributed architecture was designed. The controller was then plugged to the Aimsun micro-simulator. The simulated scenario was based on real traffic data collected from the south ring of Grenoble. Were examined both, the balancing metric (optimized) and a set of standard traffic metrics (not optimized). The results showed that the balancing has a positive impact on the traffic flow, in particular, by smoothing the vehicle dynamics, it can potentially increase the average velocity (and thus, reduce the travelling time) and reduce the fuel consumption (and related emissions). The proposed modular architecture enabled to perform the optimization for long freeway sections in the real-time.

6.6.3. Control of urban traffic networks

Participants: C. Canudas de Wit [Contact person], F. Garin, P. Grandinetti.

This work deals with efficient operation of urban traffic networks, by controlling traffic lights. A first contribution has been to devise a model for urban networks, based on the Cell-Transmission-Model adapted to signalized intersections, and then simplified with an average-based approximation. Then, based on this model, a control law has been designed, where the duty cycle of each traffic light is optimized in real time, globally considering the whole network. We have chosen a simple one-step-ahead optimization, which can be formulated as a linear program, thus resulting in simple and fast optimization. This work is part of the PhD thesis of Pietro Grandinetti.

6.6.4. Stability of Monotone Dynamical Flow Networks

Participants: E. Lovisari [Contact person], G. Como [U. Lund], K. Savla [U. of Southern California].

The stability properties of monotone dynamical flow networks are studied [22]. Demand and supply functions relate states and flows of the network, and the dynamics at junctions are subject to fixed turning rates. Our main result consists in the characterization of a stability region such that: If the inflow vector in the network lies strictly inside the stability region and a certain graph theoretical condition is satisfied, then a globally asymptotically stable equilibrium exists. In contrast, if the inflow vector lies strictly outside the region, then every trajectory grows unbounded in time. As a special case, our framework allows for the stability analysis of the Cell Transmission Model on networks with arbitrary topologies. These results extend and unify previous work by Gomes et al. on stability of the Cell Transmission Model on a line topology as well as that by the authors on throughput optimality in monotone dynamical flow networks.

6.6.5. Control of communicating vehicles in urban environment

Participants: C. Canudas de Wit [Contact person], G. de Nunzio.

The stability properties analysis of the Variable Length Model (proposed by Prof. Canudas de Wit in 2011), adapted to the urban environment was studied. It has been found that the canonical definition of Lyapunov stability for the equilibrium points does not hold for the system under analysis. A different approach for the analysis of the stability properties of the system has been introduced. Furthermore, an energy map of the equilibrium points has been obtained. Namely, a cost was associated with each feasible equilibrium point of the system, thus obtaining an assessment of the efficiency of any operation point of the system. A Variable-Speed-Limits tracking controller of the desired operation point (i.e. equilibrium) has been also devised, in order to simulate the response of the driver to the energy-efficient speed advisory. This work was submitted and accepted at the IEEE Conference on Decision and Control 2014, with the title "Urban Traffic Eco-Driving: Speed Advisory Tracking". A previous work on the steady-state analysis of the Variable Length Model in urban environment was carried out in [27]. The effort has been put also on the validation of the macroscopic model (i.e. the Variable Length Model), used for traffic evolution prediction and control synthesis. The validation procedure was run with a microscopic traffic simulator, and aims at proving that the evolution of the state of the mathematical model replicates accurately the true evolution of the traffic conditions. In particular, an important variable modeled by the system is the length of the congested area of the road section under analysis, which may be thought of as the queue length. It has been shown that the macroscopic model is able to depict the evolution in time of the queue length, with only a small error with respect to the real congestion simulated by the highly-detailed microscopic simulator. Furthermore, the validation process aims, not only at confirming the reliability of the dynamical model, but also the accuracy of the energy consumption model and the other macroscopic traffic performance metrics that have been defined in order to formulate the optimization problem. Within the COMFORT project exchange program, the work on bandwidth maximization on signalized arterials by introducing VSL as an additional degree of freedom, and by considering the energetic aspects of the problem was expanded. The canonical bandwidth maximization problem is defined as the maximization of the time interval that the vehicles can use to drive though a sequence of signalized intersections without stopping; this is achieved solely by offset control. The extension of this framework aims at showing that the additional degree of freedom (i.e. variable speed limits) improves in every case the bandwidth. A further simulation campaign in a microscopic simulator shows the benefits of the theoretical bandwidth maximization on the standard traffic performance metrics. In particular, fluidity of traffic and lower number of stops result to be highly beneficial in terms of energy consumption, without losing much in terms of traveling time.

NEUROMATHCOMP Project-Team

5. New Results

5.1. Highlights of the Year

Olivier Faugeras received the Okawa prize for his pioneering contributions for computer vision and for computational neuroscience. The ceremony will be held in Tokyo in March 2015.

5.2. Neural Networks as dynamical systems

5.2.1. Heteroclinic cycles in Hopfield networks

Participants: Pascal Chossat, Martin Krupa.

Learning or memory formation are associated with the strengthening of the synaptic connections between neurons according to a pattern reflected by the input. According to this theory a retained memory sequence is associated to a dynamic pattern of the associated neural circuit. In this work we consider a class of network neuron models, known as Hopfield networks, with a learning rule which consists of transforming an information string to a coupling pattern. Within this class of models we study dynamic patterns, known as robust heteroclinic cycles, and establish a tight connection between their existence and the structure of the coupling.

This work is available as and has been submitted to a Journal.

5.2.2. Periodic forcing of stabilized E-I networks: Nonlinear resonance curves and dynamics

Participants: Romain Veltz, Terry Sejnowski [Salk Institute].

Inhibition stabilized networks (ISNs) are neural architectures with strong positive feedback among pyramidal neurons balanced by strong negative feedback from in-hibitory interneurons, a circuit element found in the hippocampus and the primary visual cortex. In their working regime, ISNs produce damped oscillations in the γ-range in response to inputs to the inhibitory population. In order to understand the proper-ties of interconnected ISNs, we investigated periodic forcing of ISNs. We show that ISNs can be excited over a range of frequencies and derive properties of the resonance peaks. In particular, we studied the phase-locked solutions, the torus solutions and the resonance peaks. More particular, periodically forced ISNs respond with (possibly multi-stable) phase-locked activity whereas networks with sustained intrinsic oscilla-tions respond more dynamically to periodic inputs with tori. Hence, the dynamics are surprisingly rich and phase effects alone do not adequately describe the network re-sponse. This strengthens the importance of phase-amplitude coupling as opposed to phase-phase coupling in providing multiple frequencies for multiplexing and routing information.

This work has been submitted to a Journal and is available as [38].

5.3. Mean field approaches

5.3.1. A Large Deviation Principle and an Expression of the Rate Function for a Discrete Stationary Gaussian Process

Participants: Olivier Faugeras, James Maclaurin.

We prove a Large Deviation Principle for a stationary Gaussian process over $\mathbb{R}^b$, indexed by $\mathbb{Z}^d$ (for some positive integers d and b), with positive definite spectral density and provide an expression of the corresponding rate function in terms of the mean of the process and its spectral density. This result is useful in applications where such an expression is needed.

This work has been accepted for publication in Entropy [20].

5.3.2. A representation of the relative entropy with respect to a diffusion process in terms of its infinitesimal-generator

Participants: Olivier Faugeras, James Maclaurin.

In this paper we derive an integral (with respect to time) representation of the relative entropy (or Kullback-Leibler Divergence) $R(\mu|P)$, where μ and P are measures on $C([0,T];\mathbb{R}^d)$. The underlying measure P is a weak solution to a Martingale Problem with continuous coefficients. Our representation is in the form of an integral with respect to its infinitesimal generator. This representation is of use in statistical inference (particularly involving medical imaging). Since $R(\mu||P)$ governs the exponential rate of convergence of the empirical measure (according to Sanov's Theorem), this representation is also of use in the numerical and analytical investigation of finite-size effects in systems of interacting diffusions.

This work has been accepted for publication in the Journal Entropy [21].

5.3.3. Asymptotic description of stochastic networks of rate neurons with correlated synaptic weights

Participants: Olivier Faugeras, James Maclaurin.

We study the asymptotic law of a network of interacting neurons when the number of neurons becomes infinite. Given a completely connected network of neurons in which the synaptic weights are Gaussian correlated random variables, we describe the asymptotic law of the network when the number of neurons goes to infinity. Unlike previous works which made the biologically unplausible assumption that the weights were i.i.d. random variables, we assume that they are correlated. We introduce the process-level empirical measure of the trajectories of the solutions to the equations of the finite network of neurons and the averaged law (with respect to the synaptic weights) of the trajectories of the solutions to the equations of the network of neurons. The result is that the image law through the empirical measure satisfies a large deviation principle with a good rate function. We provide an analytical expression of this rate function. This work has appeared in the Comptes Rendus de l'Academie des Sciences. Serie 1, Mathematique [22].

We have continued the development, started in [22], of the asymptotic description of certain stochastic neural networks. We use the Large Deviation Principle (LDP) and the good rate function H announced there to prove that H has a unique minimum, a stationary measure on the set of trajectories $\mathcal{T}^{\mathbb{Z}}$. We characterize this measure by its two marginals, at time 0, and from time 1 to T. The second marginal is a stationary Gaussian measure. With an eye on applications, we show that its mean and covariance operator can be inductively computed. Finally we use the LDP to establish various convergence results, averaged and quenched. This work has also appeared in the Comptes Rendus de l'Academie des Sciences. Serie 1, Mathematique [23].

5.3.4. Asymptotic description of stochastic networks of integrate-and-fire neurons

Participants: François Delarue [UNS, LJAD], James Inglis [EPIs TOSCA and NeuroMathComp], S Rubenthaler [UNS, LJAD], Etienne Tanré [EPI TOSCA].

J. Inglis, together with F. Delarue (Univ. Nice – Sophia Antipolis), E. Tanré (Inria TOSCA) and S. Rubenthaler (Univ. Nice – Sophia Antipolis) completed their study of the mean-field convergence of a highly discontinuous particle system modeling the behavior of a spiking network of neurons, based on the integrate-and-fire model. Due to the highly singular nature of the system, it was convenient to work with a relatively unknown Skorohod topology. The resulting article [46] has been accepted for publication in *Stochastic Processes and Related Fields*.

5.3.5. Asymptotic description of stochastic networks of spiking neurons with dendrites

Participants: James Inglis [EPIs TOSCA and NeuroMathComp], Denis Talay [EPI TOSCA].

J. Inglis and D. Talay introduced in [49] a new model for a network of spiking neurons that attempted to address several criticisms of previously considered models. In particular the new model takes into account the role of the dendrites, and moreover includes non-homogeneous synaptic weights to describe the fact that not all neurons have the same effect on the others in the network. They were able to obtain mean-field convergence results, using new probabilistic arguments.

5.3.6. Asymptotic description of stochastic networks of realistic neurons and synapses

Participants: Mireille Bossy [EPI TOSCA], Olivier Faugeras, Denis Talay [EPI TOSCA].

In this note, we clarify the well-posedness of the limit equations to the mean-field N -neuron models proposed in [1] and we prove the associated propagation of chaos property. We also complete the modeling issue in [1] by discussing the well-posedness of the stochastic differential equations which govern the behaviour of the ion channels and the amount of available neurotransmitters.

This work has been submitted for publication to a Journal and is available as [40].

5.3.7. On the Hamiltonian structure of large deviations in stochastic hybrid systems

Participants: Paul Bressloff [Prof. University of Utah, Inria International Chair], Olivier Faugeras.

We develop the connection between large deviation theory and more applied approaches to stochastic hybrid systems by highlighting a common underlying Hamiltonian structure. A stochastic hybrid system involves the coupling between a piecewise deterministic dynamical system in $\mathbb{R}^d$ and a time-homogeneous Markov chain on some discrete space Γ. We assume that the Markov chain on Γ is ergodic, and that the discrete dynamics is much faster than the piecewise deterministic dynamics (separation of time-scales). Using the Perron-Frobenius theorem and the calculus-of-variations, we evaluate the rate function of a large deviation principle in terms of a classical action, whose Hamiltonian is given by the Perron eigenvalue of a $|\Gamma|$-dimensional linear equation. The corresponding linear operator depends on the transition rates of the Markov chain and the nonlinear functions of the piecewise deterministic system. The resulting Hamiltonian is identical to one derived using path-integrals and WKB methods. We illustrate the theory by considering the example of stochastic ion channels. This work has been submitted for publication to a Journal and is available as [41].

5.4. Neural fields theory

5.4.1. Neural fields with noise

Participants: Olivier Faugeras, James Inglis.

We extend the theory of neural fields which has been developed in a deterministic framework by considering the influence spatio-temporal noise. The outstanding problem that we here address is the development of a theory that gives rigorous meaning to stochastic neural field equations, and conditions ensuring that they are well-posed. Previous investigations in the field of computational and mathematical neuroscience have been numerical for the most part. Such questions have been considered for a long time in the theory of stochastic partial differential equations, where at least two different approaches have been developed, each having its advantages and disadvantages. It turns out that both approaches have also been used in computational and mathematical neuroscience, but with much less emphasis on the underlying theory. We present a review of two existing theories and show how they can be used to put the theory of stochastic neural fields on a rigorous footing. We also provide general conditions on the parameters of the stochastic neural field equations under which we guarantee that these equations are well-posed. In so doing we relate each approach to previous work in computational and mathematical neuroscience. We hope this will provide a reference that will pave the way for future studies (both theoretical and applied) of these equations, where basic questions of existence and uniqueness will no longer be a cause for concern. This work has appeared in the Journal of Mathematical Biology [19].

5.4.2. A center manifold result for delayed neural fields equations

Participants: Romain Veltz, Olivier Faugeras.

Lemma C.1 in [83] is wrong. This lemma is used in the proof of the existence of a smooth center manifold, Theorem 4.4. An additional assumption is required to prove this existence. We spell out this assumption, correct the proofs and show that the assumption is satisfied for a large class of delay functions τ. We also weaken the general assumptions on τ.

This work has been accepted for publication in the SIAM Journal on Mathematical Analysis and is available as [37].

5.5. Spike trains statistics

5.5.1. Exact computation of the Maximum Entropy Potential of spiking neural networks models

Participants: Bruno Cessac [correspondent], Rodrigo Cofre.

Understanding how stimuli and synaptic connectivity in uence the statistics of spike patterns in neural networks is a central question in computational neuroscience. Maximum Entropy approach has been successfully used to characterize the statistical response of simultaneously recorded spiking neurons responding to stimuli. But, in spite of good performance in terms of prediction, the fitting parameters do not explain the underlying mechanistic causes of the observed correlations. On the other hand, mathematical models of spiking neurons (neuro-mimetic models) provide a probabilistic mapping between stimulus, network architecture and spike patterns in terms of conditional probabilities. In this paper we build an exact analytical mapping between neuro-mimetic and Maximum Entropy models.

This work has been published in [18] and presented in [44], [29], [30], [27].

5.5.2. Parameter Estimation for Spatio-Temporal Maximum Entropy Distributions: Application to Neural Spike Trains

Participants: Bruno Cessac [correspondent], Hassan Nasser.

We propose a numerical method to learn maximum entropy (MaxEnt) distributions with spatio-temporal constraints from experimental spike trains. This is an extension of two papers, [10] and [4], which proposed the estimation of parameters where only spatial constraints were taken into account. The extension we propose allows one to properly handle memory effects in spike statistics, for large-sized neural networks.

This work has been published in [25] and presented in [44], [29], [30], [27].

5.6. Synaptic plasticity

5.6.1. Large Deviations of an Ergodic Synchoronous Neural Network with Learning

Participants: Olivier Faugeras, James Maclaurin.

In this work we determine a Large Deviation Principle (LDP) for a model of neurons interacting on a lattice $\mathbb{Z}^d$. The neurons are subject to correlated external noise, which is modelled as an infinite-dimensional stochastic integral. The probability law governing the noise is strictly stationary, and we are therefore able to find a LDP for the probability laws Π^n governing the ergodic empirical measure μ^n generated by the neurons in a cube of length $(2n+1)$ as n asymptotes to infinity. We use this LDP to determine an LDP for the neural network model. The connection weights between the neurons evolve according to a learning rule / neuronal plasticity, and these results are adaptable to a large variety of specific types of neural network. This LDP is of great use in the mathematical modelling of neural networks, because it allows a quantification of the likelihood of the system deviating from its limit, and also a determination of which direction the system is likely to deviate. The work is also of interest because there are nontrivial correlations between the neurons even in the asymptotic limit, thereby presenting itself as a generalisation of traditional mean-field models.

This work is available [47] and is under review in a Journal.

5.7. Visual Neuroscience

5.7.1. On the effects on cortical spontaneous activity of the symmetries of the network of pinwheels in visual area V1

Participants: Romain Veltz, Pascal Chossat, Olivier Faugeras.

We consider the problem of describing mathematically the spontaneous activity of V1 by combining several important experimental observations including 1) the organization of the visual cortex into a spatially periodic network of hypercolumns structured around pinwheels, 2) the difference between short-range and long-range intra-cortical connections, the first ones being rather isotropic and producing naturally doubly-periodic patterns by Turing mechanisms, the second one being patchy and 3) the fact that the Turing patterns spontaneously produced by the short-range connections and the network of pinwheels have similar periods. By analyzing the Preferred Orientation (PO) map, we are able to classify all possible singular points of the PO maps (the pinwheels) as having symmetries described by a small subset of the wallpaper groups. We then propose a description of the spontaneous activity of V1 using a classical voltage-based neural field model that features isotropic short-range connectivities modulated by non-isotropic long-range connectivities. A key observation is that, with only short-range connections and because the problem has full translational invariance in this case, a spontaneous doubly-periodic pattern generates a 2-torus in a suitable functional space which persists as a flow-invariant manifold under small perturbations, hence when turning on the long-range connections. Through a complete analysis of the symmetries of the resulting neural field equation and motivated by a numerical investigation of the bifurcations of their solutions, we conclude that the branches of solutions which are stable over an extended set of parameters are those corresponding to patterns with an hexagonal (or nearly hexagonal) symmetry. The question of which patterns persist when turning on the long-range connections is answered by 1) analyzing the remaining symmetries on the perturbed torus and 2) combining this information with the Poincaré-Hopf theorem. We have developed a numerical implementation of the theory that has allowed us to produce the patterns of activities predicted by the theory, the planforms. In particular we generalize the contoured and non-contoured planforms predicted by previous authors and predict the existence of mixed contoured/non-contoured planforms. We also found that these planforms are most likely to be time evolving. This work is available as a preprint [36] and has been submitted to a Journal.

5.7.2. Decoding the retina with the first wave of spikes

Participants: Geoffrey Portelli, John Barrett [Institute of Neuroscience, Medical School, Newcastle University, Newcastle UK], Evelyne Sernagor [Institute of Neuroscience, Medical School, Newcastle University, Newcastle UK], Timothée Masquelier [Institut de la Vision, UPMC Université Paris 06, Paris, France], Pierre Kornprobst [correspondent].

How a population of retinal ganglion cells (RGCs) encode the visual scene remains an open question. Several coding strategies have been investigated out of which two main views have emerged: considering RGCs as independent encoders or as synergistic encoders, i.e., when the concerted spiking in a RGC population carries more information than the sum of the information contained in the spiking of individual RGCs. Although the RGCs assumed as independent encode the main information, there is currently a growing body of evidence that considering RGCs as synergistic encoders provides complementary and more precise information. Based on salamander retina recordings, it has been suggested [66] that a code based on differential spike latencies between RGC pairs could be a powerful mechanism. Here, we have tested this hypothesis in the mammalian retina. We recorded responses to stationary gratings from 469 RGCs in 5 mouse retinas. Interestingly, we did not find any RGC pairs exhibiting clear latency correlations (presumably due to the presence of spontaneous activity), showing that individual RGC pairs do not provide sufficient information in our conditions. However considering the whole RGC population, we show that the shape of the wave of first spikes (WFS) successfully encodes for spatial cues. To quantify its coding capabilities, we performed a discrimination task and we showed that the WFS was more robust to the spontaneous firing than the absolute latencies are. We also investigated the impact of a post-processing neural layer. The recorded spikes were fed into an artificial lateral geniculate nucleus (LGN) layer. We found that the WFS is not only preserved but even refined through the LGN-like layer, while classical independent coding strategies become impaired. These findings suggest that even at the level of the retina, the WFS provides a reliable strategy to encode spatial cues.

This work is ongoing and was presented as a poster at CNS 2014. See [26] for more details.

5.7.3. Microsaccades enable efficient synchrony-based visual feature detection by retinal ganglion cells

Participants: Timothée Masquelier [Institut de la Vision, UPMC Université Paris 06, Paris, France], Geoffrey Portelli, Pierre Kornprobst [correspondent].

Fixational eye movements are common across vertebrates, yet their functional roles, if any, are debated [70]. To investigate this issue, we exposed the Virtual Retina simulator [84] to natural images, generated realistic drifts and microsaccades [59], and analyzed the output spike trains of the parvocellular retinal ganglion cells (RGC). We first computed cross-correlograms between pairs of RGC that are strongly excited by the image corresponding to the mean eye position. Not surprisingly, in the absence of eye movements, that is when analyzing the tonic (sustained) response to a static image, these cross-correlograms are flat. Adding some slow drift ($\approx$20 min/s, self- avoiding random walk) creates long timescale (>1s) correlations because both cells tend to have high firing rates for central positions. Adding microsaccades ($\approx 0.5°$ in 25ms, that is $\approx 20°$/s) creates short timescale (tens of ms) correlations: cells that are strongly excited at a particular landing location tend to spike synchronously shortly after the landing. What do the patterns of synchronous spikes represent? To investigate this issue, we fed the RGC spike trains to neurons equipped with spike timing-dependent plasticity (STDP) and lateral inhibitory connections [73]. Neurons self-organized, and each one selected a set of afferents that consistently fired synchronously. We then reconstructed the corresponding visual stimuli by convolving the synaptic weight matrices with the RGC receptive fields. In most cases, we could easily recognize what was learned (e.g. a face), and the neuron was selective (e.g. only responded for microsaccades that landed on a face). Without eye movements, or with only the drift, the STDP- based learning failed, because it needs correlations at a timescale roughly matching the STDP time constants [65]. Microsaccades are thus necessary to generate a synchrony-based coding scheme. More specifically, after each microsaccade landing, cells that are strongly excited by the image corresponding to the landing location tend to fire their first spikes synchronously. Patterns of synchronous spikes can be decoded rapidly – as soon as the first spikes are received – by downstream “coincidence detector” neurons, which do not need to know the landing times. Finally, the required connectivity to do so can spontaneously emerge with STDP. As a whole, these results suggest a new role for microsaccades – to enable efficient visual feature learning and detection thanks to synchronization – that differs from other proposals such as time-to-first spike coding with respect to microsaccade landing times.

This work is ongoing and was presented as a poster at CNS 2014.

5.7.4. A new retina-inspired descriptor for image classification

Participants: Cristina Hilario [Pattern Analysis and Computer Vision, PAVIS, Istituto Italiano di Tecnologia, Genova, Italy], Diego Sona [Pattern Analysis and Computer Vision, PAVIS, Istituto Italiano di Tecnologia, Genova, Italy], Kartheek Medathati, Pierre Kornprobst [correspondent].

Recent studies on the visual system reveal that retina is smarter than scientists believed. One low level processing occurring at the retina is feature extraction, becoming an inspiration to build novel image descriptors for image categorization. However only few methods have taken advantage of this idea, such as FREAK descriptor [51], which consists of a circular grid of a concentric distribution of overlapping receptive fields (RFs) in which average image intensities are compared pairwise. In this work we extended such a descriptor but sticking much more to biological data and models of retina. Each RF in our model is described with a linear-nonlinear model (LN) taking into account inhibitory surrounds with parameters based on biological findings. Based on the activity of retinal ganglion cells, we investigated several methods to define a set of descriptors. The performance of each descriptor was tested on computer vision datasets for texture and scene categorization.

This work is ongoing and was presented as a poster at the 1st Workshop of Visual Image Interpretation in Humans and Machine (VIIHM, EPSRC Network for Biological and Computer Vision in the UK).

5.7.5. Shifting stimulus for faster receptive estimation of ensembles of neurons

Participants: Daniela Pamplona, Bruno Cessac, Pierre Kornprobst [correspondent].

The spike triggered averaged (STA) technique has been widely used to estimate the receptive fields (RF) of sensory neurons [58]. Theoretically, it has been shown that when the neurons are stimulated with a white noise stimulus the STA is an unbiased estimator of the neuron RF (up to a multiplicative constant). The error decreases with the number of spikes at a rate proportional to the stimulus variance [75]. Experimentally, for visual neurons, the standard stimuli are checkerboards where block size is heuristically tuned. This raises difficulties when dealing with large neurons assemblies: When the block size is too small, neuron's response might be too weak, and when it is too large, one may miss RFs. Previously online updating the stimulus in the direction of larger stimulus-neural response correlation [62] or mutual information [69], [68] has been proposed. However, these approaches can not be applied for an ensemble of cells recorded simultaneously since each neuron would update the stimulus in a different direction. We propose an improved checkerboard stimulus where blocks are shifted randomly in space at fixed time steps. Theoretically, we show that the STA remains an unbiased estimator of the RF. Additionally, we show two major properties of this new stimulus: (i) For a fixed block sized, RF spatial resolution is improved as a function of the number of possible shifts; (ii) Targeting a given RF spatial resolution, our method converges faster than the standard one. Numerically, we perform an exhaustive analysis of the performance of the approach based on simulated spiked trains from LNP cascades neurons varying RF sizes and positions. Results show global improvements in the RF representation even after short stimulation times. This makes this approach a promising solution to improve RF estimation of large ensemble of neurons.

This work is ongoing and was submitted to COSYNE 2015.

5.7.6. Shifting stimulus for faster receptive estimation of ensembles of neurons

Participants: Kartheek Medathati, Fabio Solari [University of Genoa - DIBRIS, Italy], Manuela Chessa [University of Genoa - DIBRIS, Italy], Guillaume S. Masson [Institut des Neurosciences de la Timone, Team InVibe], Pierre Kornprobst [correspondent].

Motion estimation has been studied extensively in neurosciences in the last two decades. The general consensus that has evolved from the studies in the primate vision is that it is done in a two stage process involving cortical areas V1 and MT in the brain. Spatio temporal filters are leading contenders in terms of models that capture the characteristics exhibited in these areas. Even though there are many models in the biological vision literature covering the optical flow estimation problem based on the spatio-temporal filters little is known in terms of their performance on the modern day computer vision datasets such as Middlebury. In this paper, we start from a mostly classical feedforward V1-MT model introducing a additional decoding step to obtain an optical flow estimation. Two extensions are also discussed using nonlinear filtering of the MT response for a better handling of discontinuities. One essential contribution of this paper is to show how a neural model can be adapted to deal with real sequences and it is here for the first time that such a neural model is benchmarked on the modern computer vision dataset Middlebury. Results are promising and suggest several possible improvements.

This work is ongoing and was presented as a poster at the 1st Workshop of Visual Image Interpretation in Humans and Machine (VIIHM, EPSRC Network for Biological and Computer Vision in the UK). See [35] for more details.

5.7.7. Exploring the richness of center-surround dynamics: A bifurcation study

Participants: Kartheek Medathati, James Rankin [Center for Neural Sciences, NYU, USA], Guillaume S. Masson [Institut des Neurosciences de la Timone, Team InVibe], Pierre Kornprobst [correspondent].

The balance of excitatory and inhibitory interactions between neurons is one of the characteristic aspects of neural computation. In both neural network and neural field models these interactions have been modeled using center-surround connectivity kernels. Depending on the relative strength of excitation and inhibition these networks have been found to exhibit rich and interesting dynamical behavior. Although many models have been reported in the literature using center-surround connectivity kernels and many experimental studies have shown evidence for changes in observed behavior from winner-take-all to gain control, a thorough bifurcation analysis of these networks in terms of sensitivity of the network to peak strength, discriminability of the peaks and speed of convergence has not been done. In our present work we visit this question in order to identify the parameter regimes where this important switch in the behavior of the network occurs and also establish the trade offs that arise with the choice of a particular connectivity kernel.

This work is ongoing and was presented as a poster at the conference "Nonlinear dynamics and stochastic methods: from neuroscience to other biological applications"

5.7.8. From Habitat to Retina:Neural Population Coding using Natural Movies

Participants: Bruno Cessac [correspondent], Ruben Herzog [Centro Interdisciplinario de Neurociencia de Valparaíso, Univ de Valparaíso, 2360102 Valparaíso, Chile], Joaquin Araya [Centro Interdisciplinario de Neurociencia de Valparaíso, Univ de Valparaíso, 2360102 Valparaíso, Chile], Michael Pizarro [Centro Interdisciplinario de Neurociencia de Valparaíso, Univ de Valparaíso, 2360102 Valparaíso, Chile], Cesar Ravello [Centro Interdisciplinario de Neurociencia de Valparaíso, Univ de Valparaíso, 2360102 Valparaíso, Chile], Maria Jose Escobar [Universidad Técnica Federico Santa María, Valparaiso, Chile], Adrian Palacios [Centro Interdisciplinario de Neurociencia de Valparaíso, Univ de Valparaíso, 2360102 Valparaíso, Chile].

We use a diurnal rodent retina (O. Degus), which has the advantage of present a 3:1 proportion of rods and cones, respectively, to study the RGC population responses to habitat-based natural stimuli. In order to do this, we have developed a mobile robot that is capable to record movies in the natural habitat of this rodent, simulating both his movements and the eye-ground distance, which allows us to stimulate and record an in vitro retina patch using MEA (multi electrode array) with a sequence of images taken from the animal natural habitat. The analysis of spike statistics has been done using the Enas software to characterize spatio-temporal pairwise correlation with Gibbs distributions. potential constitutes a useful tool for comparing pairwise spatio-temporal correlations between different conditions for the same RGC population. We show that correlated spiking patterns represents a major deviation between White Noise and Natural Movies conditions. We also conclude that population coding for this monophasic OFF RGC population is mostly based on spatial correlation when stimulated with Natural Movies.

This ongoing work has been presented in [48].

...

NEUROSYS Team

6. New Results

6.1. Highlights of the Year

Microscopic action affects mesoscopic and macroscopic action in neural systems. In the context of general anaesthesia, it is not understood how single neuron properties, such as ion-channel conductivities or anesthetic action on neuron receptors, translate to population dynamics and consequently to behavior. The work of Laure Buhry and Axel Hutt [4] proposes a modelling approach how to bridge the microscopic and the mesoscopic scale. The most interesting aspect is that this model bridge allows to extend standard neural field theory on the mesoscopic scale instead of introducing a new model.
In addition, we have developed strong collaborations with medical doctors. First, we have established a collaboration with Dr. Denis Schmartz and Dr. Claude Meistelmann at the *CHU Nancy* to plan and perform well-controlled resting state experiments under propofol anaesthesia. Second, we are in close contact to Jean-Luc Schaff at the *CHU Nancy* (together with Laurent Koessler at *CRAN*) in the context of sleep monitoring. Dr. Schaff has provided us polysomnographic data measured during sleep of insomnia patients.

6.2. From the microscopic to the mesoscopic scale

Participants: Laure Buhry, Axel Hutt, Francesco Giovannini, LieJune Shiau
The Highlight of the Year bridges the microscopic scale and the mesoscopic scale. One partial result has already been used in one of our publications [3] to study the link between population dynamics on the mesoscopic scale and the EEG on the macroscopic scale.

In addition, the work of Francesco Giovannini aims at gaining a better understanding of the effects of anaesthesia on the neural correlates of memory, focusing on how anaesthetics disrupt the interaction between the hippocampus and the cerebral cortex. Studies have shown that these two brain structures exhibit a strong synchronisation of their respective neural activity, when performing memory tasks. Neurophysiology experiments have identified various possible candidate generators for rhythmic activity in the area CA1, CA3 and Dentate Gyrus areas of the hippocampus. However the mechanisms by which cortico- hippocampal synchronisation is elicited, and maintained, are yet to be fully understood. As a first step towards this objective, Francesco obtained a working mathematical model of a biologically plausible hippocampal CA1-3 neural cell, based on the Hodgkin-Huxley neuron, capable of exhibiting long-lasting persistent firing activity when subject to a strong transient stimulus. This behaviour is underlay by an intrinsic membrane current activated by the increase of intracellular Calcium ions, following the discharge of an action potential by the neuron. Our hypothesis is that large ensembles of such persistent-firing neurons could sustain the memory-related rhythmic activity displayed by the hippocampus. In this context, Laure Buhry and Axel Hutt work with LieJune Shiau (University of Houston) on a better understanding of the models used by the community of computational neuroscientists. The goal is to show in which extent models are comparable or interchangeable. We focus on the comparison of oscillatory mechanisms of neuronal populations in different spiking models, especially in the Hodgkin-Huxley and the adaptive exponential integrate-and-fire model.

These latter studies link the two description scales by a bottom-up approach.

Conversely, Axel Hutt and collaboration partners from the University of Noth Carolina - Chapel Hill have analysed Local Field Potentials measured in ferrets prefrontal cortex and visual cortex under anesthesia in a top-down analysis [21]. This data allows to extract network interactions in prefrontal cortex and visual cortex and hence revealing underlying mechanisms in general anaesthesia.

6.3. From the mesoscopic to the macroscopic scale

Participants: Laurent Bougrain, Axel Hutt, Pedro Garcia-Rodriguez, Eric Nichols, Guillaume Serrière, Tamara Tosic, Nicole Voges, Mariia Fedotenkova, Meysam Hashemi, Cecilia Lindig-Leon, Kevin Green, Sébastian Rimbert, Thomas Tassone.
To understand the action of anaesthetic drugs on the EEG-signal observed experimentally, Meysam Hashemi has developed and studied several neural mass models [18], [15], [16], [3]. He has identified the thalamo-cortical loop (TCL) as a possible origin of $\delta-$activity. Since loss of consciousness is accompanied by emerging $\delta-$activity, this work relates the TCL to the loss of consciousness.

Increasing the anaesthetic concentration beyond the point of loss of consciousness, EEG-signals exhibit alternating patterns of high and low activity. This activity is called burst suppression. Since these alternations resemble stochastic jumps between low and high activity resting states, Pedro Garcia-Rodriguez and colleagues are working on a stochastic theory based on neural mass models to describe and reproduce these experimental results. Since the minimum mathematical model for such an effect is two-dimensional and does not exhibit potential dynamics, whereas the majority of literature up to date considers one-dimensional stochastic models obeying potential dynamics, Pedro and colleagues had to develop a new stochastic theory. They can show that the two-dimensional dynamics of the neural mass model can be mapped to a one-dimensional stochastic potential model [14], [13]. This reduction allows to apply standard stochastic theory to describe burst suppression as stochastic transistions. This finding indicates the presence of multiple resting states in the brain and supports a heavily discussed hypothesis on the loss of consciousness.

Biological neural networks are subject to random fluctuations, originating from intrinsic random fluctuations of ions or from external stimulus. The latter neural mass models take into account these fluctuations by assuming additive random input fluctuations. For many decades, these additive fluctuations have been assumed to not affect the stability of the system. However, previous own work has revealed that additive fluctuations tune the stability of nonlinear high-dimensional systems. Since random fluctuations play an important role in the description of neural population dynamics and realistic models consider , it is necessary to study in detail how random fluctuations affect the stability of neural mass models and, hence, how our mathematical model analyses have to be modified. To this end, Axel Hutt and colleagues have performed a stochastic center manifold analysis in a delayed stochastic neural mass model [5] and have found conditions for the stability shift. A first application to delayed stochastic neural fields has revealed how additive random fluctuations may affect EEG-signals [19], [6], however additional detailed mathematical studies and the comparison to experimental data are necessary to affirm the importance of the stochastic effect. Essentially, this work emphasizes to take into account nonlinear noise effects in neural mass and neural field models.

Neural mass models do not consider the spatial extension of neural populations and consequently neglect transmission or interaction delay between neurons at different spatial locations. Taking into account the spatial extension and axonal transmission delay, Axel Hutt and colleagues have shown mathematically [7] how travelling activity fronts propagate through neural tissue and how the fronts properties, such as speed, depend on the neural field properties.

The latter neural field model is embedded in a one-dimensional space. Since biological neural populations in the neocortex are organized in two-dimensional layers or sheets, it is necessary to employ neural field models in two spatial dimensions. This causes both theoretically and numerically problems in the presence of axonal transmission delay. Eric Nichols and colleagues has implemented a recent numerical integration algorithm [8] in the visualization software NeuralFieldSimulator, cf. section5.1 . This software is the basis of numerical bifurcation studies of two-dimensional neural field models [12], [20]. First analytical results [10] show good accordance to numerical results obtained by the NeuralFieldSimulator.

The latter neural field models assume homogeneous spatial interactions, i.e., neural interactions whose strength just depends on the distance between the two neurons. This assumption is strong and not biologically realistic in certain brain areas. In addition, this assumption constrains the model description of recurrent sequences of EEG patterns, which have been found experimentally, e.g., during the emergence from general anaesthesia. Consequently to be able to describe such recurrent EEG-pattern sequences, it is necessary to improve the mathematical description of EEG-patterns. A promising new model has been derived by Axel Hutt and collagues based on heterogeneous neural fields [1]. In order to extract the recurrence EEG-patterns from data, we have extended a recent recurrence analysis technique [2]. The next step will consist in the combination of the heterogeneous neural field model and the results from the recurrence analysis.

Recurrence analysis extracts temporally reccurrent time windows in multi-dimensional datasets. Typical EEG-signals obtained durin surgery under anaesthesia include one electrode and hence a single time series only. To extract recurrence structures of such one-dimensional signals, Mariia Fedotenkova computes the multi-dimensional time-frequency representation of the signal and has worked out the best analysis technique for this step [9]. In the next step she will compute the recurrence plot for a large dataset of 110 patients under surgery (data obtained from University of Auckland).

In order to understand immobility during anaesthesia and how to supervise unconscious patients automatically in hospital emergency rooms, Cecilia Lindig-Leon studies motor imagery and its detection by BCI techniques. Limb movement execution or imagination induce sensorimotor rhythms that can be detected in EEG recordings. Her recent work considers signal power changes in two frequency bands to detect the elicited EEG rebound, i.e. the increasing of synchronization, at the end of motor imageries. The analysis is based on the database 2a of the BCI competition IV and shows that rebound can be stronger over the alpha frequency band (8-12Hz) than the beta frequency band (12-20Hz). She can demonstrate that the analysis of the alpha frequency band improves the detection of the end of motor imageries. In this context, Cecilia has compared intrinsic multi-class classifiers (i.e., one-step methods) with ensembles of two-class classifiers on dataset 2a of the BCI competition IV for motor imagery. Subsequently, she has compared the classical Common Spatial Pattern (CSP) approach and the CSP by Joint Approximate Diagonalization in order to identify whether the latter method represents an outperforming alternative.

Sleep is strongly related to anaesthesia and we have started working on the improvement of sleep monitors. The basic idea is to consider not only EEG-signals but multiple different physiological signals (e.g. heart pulses, electrocardiogram, EEG, respiration cycle, body movements) to classify sleep stages. By virtue of the different signal natures of different physiological signals, it is challenging to put together these so-called multi-modal signals in a single analysis method. To this end, Tamara Tosic and colleagues employ recurrence analysis techniques which allow to estimate time windows exhibiting temporal synchronization between physiological signals [11]. They have developed a method that is based on artificial data sets and Local Field Potentials measured under anaesthesia. In the next step, applications to sleep data (obtained from CHU Nancy) will allow to extract sleep stages and will evaluate the method.

NON-A Project-Team

6. New Results

6.1. Highlights of the Year

- We are becoming world-recognized on homogeneous approach to estimation and control [13], [24].
- New method of stability analysis and control design for time-delay systems: Implicit Lyapunov-Krasovski Functionals [72].
- New dynamical model of population of oysters for water quality monitoring [44].
- New local path planning algorithm for mobile robots based on intermediate objectives [33].
- New patent on method and device for detecting a failure on an aircraft [85].
- New book on robust control design [82].

6.2. Homogeneity Theory And Analysis Of Nonlinear Systems

Homogeneity is a kind of symmetry, if it is presented in a system model, then it may simplify analysis of stability and performance properties of the system. The new results obtained in 2014 are as follows:

- The problem of scalability of trajectories in homogeneous and locally homogeneous systems is considered [46]. It is shown that the homogeneous systems have scalability property, and locally homogeneous systems possess this property approximately.
- Constructive conditions for verification of input-to-state stability property for discontinuous systems using geometric homogeneity have been proposed in [48]. The characterization of the asymptotic gain for such systems has been presented in [47].
- The problem of finite-time output stabilization of the double integrator is addressed in [14] applying the homogeneity approach. Robustness and effects of discretization on the obtained closed loop system are analyzed.
- The paper [24] extends notion of homogeneity to the time-delay nonlinear systems. Generalizations and specifications of the homogeneity approach to time-delay nonlinear systems are given in [57], where, for instance, the stability independently on delay has been analyzed.
- In [75] the uniform stability notion for a class of non-linear time-varying systems is studied using the homogeneity framework. The results are applied to the problem of adaptive estimation for a linear system.
- The Implicit Lyapunov Function (ILF) method has been applied for homogeneous differentiator design [70]. The procedure for adjustment of differentiator parameters has been resented in the form of semi-definite programming problem. ILF-based algorithms of robust finite-time and fixed-time stabilization of the chain of integrators were developed in [34]. In [69] they were adapted for the second order sliding mode control design.
- The tutorial on homogeneous methods in high sliding mode control has been published [13]. It stresses some recently obtained results of the team about homogeneity for differential inclusions and robustness with respect to perturbations in the context of input-to-state stability.

6.3. Model-Free Control

The model free control techniques form a new and quickly developing area of control theory. It has been established by the team members and nowadays these tools find many practical applications and attract a lot of attention due to their clear advantages for designers: they provide a control law independently in the model knowledge. The achievements obtained in 2014 are as follows:

- The paper [67] proposes a motion planning approach for non-holonomic mobile robots using i-PID controller. The effectiveness and the robustness of the proposed method are shown via several simulations.
- In [60] we show that the open-loop transfer functions and the stability margins may be defined within the recent model-free control setting. Several convincing computer experiments are presented including one which studies the robustness with respect to delays.

6.4. Algebraic Technique For Estimation, Differentiation And Its Applications

Elementary techniques from operational calculus, differential algebra, and non-commutative algebra lead to a new algebraic approach for estimation and detection. It is investigated in various areas of applied sciences and engineering. The following lists only some applications:

- The article [19] presents an algebraic on-line parameters estimation method for Linear Time Invariant (LTI) systems subject to polynomial perturbations. Particular attention is given to practical implementation.
- In the paper [43], we extend the modulating functions method to estimate the state and the unknown input of a linear time-varying system defined by a linear differential equation. Numerical results are given to show the accuracy and the robustness of the proposed estimators against corrupting noises.
- The paper [17] presents a new sensorless parameter identification method for permanent magnet stepper motors. The method exploits derived linear parameterizations and least-squares algorithms.
- In [36] a continuous-time least-squares parameter estimation method through evolution equations is proposed. A deterministic framework for the estimation under noisy measurements is proposed using a Sobolev space with negative index to model of the noise.
- Causation between time series is a most important topic in econometrics, financial engineering, biological and psychological sciences, and many other fields. A new setting is introduced in [42] for examining this rather abstract concept. The corresponding calculations, which are much easier than those required by the celebrated Granger-causality, do not necessitate any deterministic or probabilistic modeling.
- The paper [59] proposes a solution to the problem of velocity and position estimation for a class of oscillating systems whose position, velocity and acceleration are zero mean signals. The proposed scheme considers that the dynamic model of the system is unknown and only noisy acceleration measurements are available.
- The communications [63], [78] are devoted to solar irradiance and irradiation short-term forecasts, which are useful for electricity production. Several different time series approaches are employed.
- In [68] we present a simple algorithm to compute the factors of a Unimodular-Upper polynomial matrix decomposition. Such decomposition is useful for spatial multiplexing in multi-input multi-output(MIMO) channel transmission system since it enables to reduce the MIMO channel matrix into independent channels by a pre- and post-filtering.
- A fault-tolerant control method based on algebraic derivative estimation is introduced in [32]. It is applied on an electromagnetically supported plate as an example of a nonlinear and an open-loop unstable system.

6.5. Observability And Observer Design For Nonlinear Systems

Observability analysis and observer design are important issues in the field of control theory. Some recent results are listed below:

- The paper [12] deals with the observability analysis of linear time systems whose outputs are affected by unknown inputs. Three different definitions of observability are proposed. Sufficient conditions are deduced for each proposed definition.
- In [11] a method of the state estimation is proposed for a class of nonlinear systems with unknown inputs whose dynamics is governed by differential-algebraic equations (DAE). The estimation is done using a sliding mode high order differentiator.
- The recent algebraic parametric method proposed by Fliess and Sira-Ramirez has been extended to numerical differentiation problem in noisy environment [66]. The obtained algebraic differentiators are non-asymptotic and robust against corrupting noises.
- The paper [41] investigates the observer design problem of nonlinear impulsive systems with impact perturbation. By using the concept of normal form, it proposes a full order finite time observer, which guarantees the finite time convergence independent of the impact perturbation.
- The development of adaptive observer techniques for nonlinear systems in the output canonical form is proposed in [22] applying additional impulsive feedback in the observer equations. The stability is investigated.
- In [55] the problem of adaptive observer design in the presence of disturbances is studied, and an augmented adaptive observer is proposed using sliding mode methodology.

6.6. Sliding Mode Control And Estimation

Sliding mode algorithms are very popular for finite-time estimation and regulation. The recent results obtained by the group are as follows:

- In [71] the high-order sliding mode control design algorithm has been developed for MIMO system using ILF Method. Procedure for tuning of control parameters is presented using Linear Matrix Inequalities.
- A novel hybrid automaton admitting the modeling of both conventional and modern(high order) sliding mode systems is presented [65]. A scheme for defining hybrid-automaton executions beyond Zeno points is proposed by means of introduction of Filippov-like executions.
- The paper [35] surveys mathematical tools required for stability/convergence analysis of modern sliding mode control systems and introduces the generalized Lyapunov theorems. Application of these results to finite-time stability analysis and settling time estimation of twisting second order sliding mode controller are given [73].
- The problem of the sliding mode control design is considered in [81] for the linear time-invariant disturbed system with the noised measurements of the output. The control law, which provides to the closed-loop system the optimal reaching (as close as possible) of the selected sliding surface, is designed using minimax state observer.
- The paper [50] deals with a signal-based method for robust and early detection of lock-in-place failures (a.k.a. jamming) in aircraft control surface servo-loops. The signal-based scheme is proposed using a sliding-mode differentiator. The developed monitoring scheme has been tested on Airbus test facilities located at Toulouse, France.
- In the paper [79], we investigate the problem of adaptive observer for simultaneous estimation of state and parameter for a class of nonlinear systems.Necessary condition for the existence of such an observer is derived. The paper [76] uses developed technique for states estimation and parameter identification for nonlinear Dengue epidemic model.
- The paper [80] investigates the problem of global finite-time observer design for a class of nonlinear systems which can be transformed into the output depending normal form.

6.7. Non-Linear, Sampled And Time-Delay Systems

Nonlinearities, sampling, quantization and time-delays cause serious obstructions for control and observer design in many fields of techniques and engineering (e.g. networked and internet systems, distributed systems etc.). The proposed by the team algebraic approach suits well for estimation and regulation in such a type of systems. The recent results are listed below:

- The method of Implicit Lyapunov-Krasovski Functional for stability analysis of time-delay systems is introduced in [72].
- The article [31] proposes a convex optimization approach for the design of relay feedback controllers. Furthermore, the approach is used in the sampled-data case in order to guarantee (locally) the practical stabilization to a bounded ellipsoid of the order of the sampling interval.
- The paper [40] addresses the controller design problem for bilateral teleoperation over unreliable networks. The stability and tracking performance analysis are presented for a novel force-reflecting emulator control scheme.
- The problem of time optimal control design is considered for a chain of integrators in [74]. The suboptimal continuous ILF-based solution is presented and compared with the optimal discontinuous feedback.
- In the erratum [26] recently proposed conditions on finite-time stability in time-delay systems are revisited and it is shown that they are incorrect. General comments on possibility of finite-time convergence in time-delay systems and a necessary condition are given.
- The problem of formulation of an equivalent characterization for instability is considered in [56]. The necessary part of the Chetaev's theorem on instability is formulated. Using the developed necessary instability conditions, the Anti-control Lyapunov Function (ALF) framework is extended and the Control Chetaev Function (CCF) concept is proposed as a counterpart of the Control Lyapunov function (CLF) theory.
- The paper [25] extends the notion of oscillations in the sense of Yakubovich to hybrid dynamics. Several sufficient stability and instability conditions for a forward invariant set are presented. The consideration is illustrated by analysis of a model of two-link compass-gait biped robot.
- The paper [15] deals with the design of an active fault-tolerant control strategy based on the supervisory control approach technique for linear time invariant MIMO systems affected by disturbances, measurement noise, and faults.
- The problem of phase regulation for a population of oscillating systems is considered in [21]. The proposed control strategy is based on a Phase Response Curve (PRC) model of an oscillator.
- The paper [51] deals with the design of an estimator-based supervisory Fault Tolerant Control scheme for Linear Time Invariant systems. A formal stability proof based on dwell-time conditions is established.
- In [39], we propose a general statistical framework for model based compressive sensing, where both sparsity and structure priors are considered simultaneously. It is based on the Latent Variable Analysis and the Gamma-Gaussian modelling.
- The paper [64] investigates the left invertibility for nonlinear time delay system with internal dynamics under some assumptions imposed on the internal dynamics. Causal and non causal estimation of the unknown inputs are respectively discussed, and the high-order sliding mode observer is used to estimate the observable states.
- In the paper [54] a simple second order model is proposed for modeling the pressure dynamics with a pure time delay on the control input. The Artstein transformation is applied in order to design the stabilizing robust nonlinear controller.

6.8. Set-Theoretic Methods of Control And Estimation

In many cases due to parametric and/or signal uncertainties presented in a plant model it is not possible to design a conventional observer, which provides a point-wise estimate of state in a finite time or asymptotically. In this case it is still frequently possible to design observers, which generate an estimate on the set of the admissible values of the state at the current instant of time. The recent new results in this field are listed below:

- An interval observer for Linear Time-Varying systems is proposed in [38]. A constructive approach to obtain a time-varying change of coordinates, ensuring the cooperativity of the observer error in the new coordinates, is provided in order to simplify the design of the interval observer.
- In [58] the problem of interval observer design is addressed for a class of descriptor linear systems with delays. An interval observation for any input in the system is provided. The control input is designed together with the observer gains in order to guarantee interval estimation and stabilization simultaneously.
- The estimation problem of a system with unknown time-delay and unknown input gains is considered in [49]. The interval observation technique is applied in order to obtain guaranteed interval of the system state.
- The book [82] introduces newly developed robust control design technique for a wide class of continuous-time dynamical systems called the “attractive ellipsoid method.” It studies nonlinear affine control systems in the presence of uncertainty and presents a constructive and easily implementable control strategy that guarantees certain stability properties.

6.9. Networked Robots

The mobile robots constitute an important area of practical development for the team:

- The paper [33] presents a path planning algorithm for autonomous navigation of non-holonomic mobile robots in complex environments. The irregular contour of obstacles is represented by segments. The optimal path planning problem is formulated as a constrained receding horizon planning problem and the trajectory is obtained by solving an optimal control problem with constraints.
- In [62] robot dynamic parameters are estimated based on power model associated with modulating functions, which avoids measuring or calculating the joint acceleration. At the same time, an advanced causal Jacobi derivative estimator is applied in order to get on-line robust derivatives from noisy measurements.
- The paper [61] provides a solution for the stabilization of a nonholonomic wheeled mobile robot which is affected by additive input disturbances. The solution is based on the supervisory control framework, finite-time stability and robust multi-output regulation.
- The demo video with the developments of NON-A team in networked robotics is given by https://www.youtube.com/watch?v=Mq_hB0UkzkY

6.10. Applications

As it was mentioned, Non-A is a kind of "method-driven" project, which deals with different aspects of finite-time estimation and control. Thus different applications are possible, ones touched this year are as follows (skipping the networked robots considered in the previous section):

- Method and device for detecting a failure on an aircraft are developed and patented [85].
- In [44] the measurements of valve activity in a population of bivalves under natural environmental conditions (16 oysters in the Bay of Arcachon, France) are used for a physiological model identification. A nonlinear auto-regressive exogenous (NARX) model is designed and tested. The developed dynamical model can be used for estimation of the normal physiological rhythms of permanently immersed oysters and, in particular, for ecological monitoring.

- The articles [53], [18], [20], [77] present novel control strategies for Permanent Magnet Synchronous Motor (PMSM), which does not ignores the relay nature of the actuators. A design procedure based on Linear Matrix Inequalities (LMI) allows us to derive the switching surfaces, which depend on the motor position. The sliding mode and nonlinear adaptive observers are designed for state estimation and parameters identification.
- The problem of air-fuel ratio stabilization in spark ignition engines is addressed in the paper [23]. The proposed strategy consists of proper switching among two control laws. The first one is based on an a priori off-line identified engine model and the second control law is adaptive. The supervisor realizes a switching rule between them providing better performance. Results of implementation on two vehicles are reported and discussed.
- The paper [37] deals with a control design for serial multicellular choppers. The novel scheme that uses two Petri nets (PNs) to carry out the control action is introduced. Experimental results from four and five-level choppers are used to emphasize the performance and the effectiveness of the proposed control scheme.
- The paper [52] is concerned with preliminary results on robot vibratory modes on-line identification using the external measurement provided by a laser tracker. A comparison between the algebraic method and the sliding modes for the parameter identification is proposed. Experimental identifications are proposed on a 6 degrees of freedom (DOF) manipulator robot Stäubli RX-170B.
- The papers [30], [29], [16] develop different fault detection schemes for robust and early detection of faults in aircraft control surfaces servo-loop. A complete Monte Carlo campaign from a high representative simulator, provided by Airbus as a part of the ADDSAFE project, as well as experimental results obtained on AIRBUS test facilities demonstrate the high fault detection performance, robustness and viability of the proposed techniques.
- The paper [28] deals with the problem of the practical tracking control of an experimental car-like system called the Robucar - a four-wheeled car in a single steering mode. A practical tracking controller is designed using the second-order sliding mode control. Experimental tests are presented and compared with the conventional sliding controller.
- Power converters are a very important for the control of high power systems. In the article [45] we propose a control strategy for minimizing the no-load conduction losses and analyze the transient behavior in case of load steps including output short-circuit.

NUMED Project-Team

5. New Results

5.1. Highlights of the Year

Vincent Calvez has been awareded an ERC Grant and the prestigious Bronze medal CNRS.

OAK Project-Team

6. New Results

6.1. Highlights of the Year

The year has allowed reaching important results in four research areas of the group: query-based why-not provenance with explanations , minimal query reformulations under constraints [14], Linked Open Data analytics , and RDF data management in the cloud .
BEST PAPERS AWARDS :
[6] **Extending Database Technology (EDBT)**. N. BIDOIT, M. HERSCHEL, K. TZOMPANAKI.
[10] **23rd International World Wide Web Conference**. D. COLAZZO, F. GOASDOUÉ, I. MANOLESCU, A. ROATIS.
[], [23] **The International Journal on Very Large Databases**. Z. KAOUDI, I. MANOLESCU.

6.2. Scalable and Expressive Techniques for the Semantic Web

A main scientific topic of the team is the design of expressive and efficient tools for analyzing and manipulating Semantic Web data, in particular RDF. Our 2014 results in this area follow three complementary directions.

First, we have finalized our model for RDF analytics and proposed a full framework in which we fully redesign, from the bottom up, core data analytics concepts and tools in the context of RDF data, leading to the first complete formal framework for warehouse-style RDF analytics. Notably, we defined $i)$*analytical schemas* tailored to heterogeneous, semantics-rich RDF graph, (ii)*analytical queries* which (beyond relational cubes) allow flexible querying of the data and the schema as well as powerful aggregation and (iii)*OLAP-style operations*. We implemented our RDF analytics platform on top of the KDB system and ported it on Postgres as well [10], [28]; work is ongoing to adapt it on a massively parallel RDF query evaluation platform, namely CliqueSquare (see below). In [24], we describe novel techniques for optimizing the evaluation of RDF analytical queries based on previously computed analytical query results.

Second, we continued our work on efficient evaluation of queries on RDF data, in the presence of constraints. *Reformulation-based query answering* is a query processing technique aiming at answering queries against data, under constraints. It consists of reformulating the query based on the constraints, so that evaluating the reformulated query directly against the data (i.e. without considering any more the constraints) produces the correct answer set. We have show how to optimize reformulation-based query answering in the setting of *ontology-based data access*, where SPARQL conjunctive queries are posed against RDF facts on which constraints expressed by an RDF Schema hold. The literature provides solutions for various fragments of RDF, aiming at computing the equivalent union of maximally-contained conjunctive queries w.r.t. the constraints. However, in general, such a union is large, thus it cannot be efficiently processed by a query engine. In this context, we have shown that generalizing the query reformulation language allows considering a space of reformulated queries (instead of a single possible choice), and selecting tthe reformulated query with lower estimated evaluation cost. We have shown experimentally that our technique enables reformulation-based query answering where the state-of-the-art approaches are simply unfeasible, while it may decrease their costs by orders of magnitude in other cases [20], [26].

Third, we have continued our work on cloud-based RDF data management. In [22], we have demonstrated CliqueSquare, a platform we developed in the team for the massively parallel processing of RDF queries. CliqueSquare enjoys the benefits of a query optimization algorithm which creates query plans as flat as possible, which in turn translates into massive opportunities for parallel processing. In [23], we have finalized our work on managing RDF data within the Amazon Web Services cloud. Finally, we have conducted a study of the existing models and algorithms published so far for the massively parrallel processing of RDF queries, which appeared as a survey in the VLDB Journal [] and was also the basis of a tutorial at the ACM SIGMOD conference.

6.3. Massively Distributed Data Management Systems

Work in this area concerning the massively parallel processing of Semantic Web data was covered within the respective module.

We have finalized our work on massively parallel processing of XML queries based on the Apache Flink framework, formerly known as Stratosphere from the Technical University of Berlin, which implements the PACT model (an expressive extension of MapReduce). In [21], we have addressed the problem of efficiently parallelizing the execution of complex nested data processing, expressed in XQuery. We provided novel algorithms showing how to translate such queries into PACT, a recent framework generalizing MapReduce in particular by supporting many-input tasks. We presented the first formal translation of complex XQuery algebraic expressions into PACT plans, and demonstrated experimentally the efficiency and scalability of our approach. The work has recently been accepted for publication to IEEE TKDE (to appear in 2015),

Finally, we have considered improving the performance of massively parallel data processing programs expressed using the PigLatin language. PigLatin is a popular language within the data management community interested in the efficient parallel processing of large data volumes. The dataflow-style primitives of PigLatin provide an intuitive way for users to write complex analytical queries, which are in turn compiled into MapReduce jobs. Currently, subexpressions occurring repeatedly in PigLatin scripts are executed as many times as they occur, leading to avoidable MapReduce jobs. The current PigLatin optimizer is not capable of recognizing, and thus optimizing, such repeated subexpressions. In [18], we have presented We present a novel approach for identifying and reusing common subexpressions occurring in PigLatin scripts. In particular, we lay the foundation of our reuse-based algorithms by formalizing the semantics of the PigLatin query language with extended nested relational algebra for bags. Our algorithm, named PigReuse, operates on the algebraic representations of PigLatin scripts, identifies subexpression merging opportunities, selects the best ones to execute based on a cost function, and merges other equivalent expressions to share its result; our experiments have confirmed the efficiency and effectiveness of our reuse-based algorithms and optimization strategies.

6.4. Advanced Algorithms for Data Querying and Transformation

We revisit in [14] the Chase&Backchase (C&B) algorithm for query reformulation under constraints. For an important class of queries and constraints, C&B has been shown to be complete, i.e. guaranteed to find all (join-)minimal reformulations under constraints. C&B is based on constructing a canonical rewriting candidate called a universal plan, then inspecting its exponentially many sub-queries in search for minimal reformulations, essentially removing redundant joins in all possible ways. This inspection involves chasing the subquery. Because of the resulting exponentially many chases, the conventional wisdom has held that completeness is a concept of mainly theoretical interest. We show that completeness can be preserved at practically relevant cost by introducing a novel reformulation algorithm that instruments the chase to maintain provenance information connecting the joins added during the chase to the universal plan subqueries responsible for adding these joins. This allows it to directly “read off” the minimal reformulations from the result of a single chase of the universal plan, saving exponentially many chases of its subqueries. We exhibit natural scenarios yielding speedups of over two orders of magnitude between the execution of the best view-based rewriting found by a commercial query optimizer and that of the best rewriting found by our algorithm.

Different types of explanations that serve as Why-Not answers have been proposed in the past and are either based on the available data, the query tree, or both. A first approach to this so called why-not provenance has been recently proposed. In [6], we show that this first approach has some shortcomings. To overcome these shortcomings, we propose Ned, an algorithm to explain data missing from a query result. NedExplain computes the why-not provenance for monotone relational queries with aggregation. This work contributes to providing necessary formalization in which the new algorithm is build. It also develops a comparative evaluation showing that it is both more efficient and effective than the state-of-the-art approach.

Solutions to answering Why-Not questions are generally more efficient and easier to interpret by developers than solutions solely based on data. However, algorithms producing such query-based explanations including ours ([6]) so far may return different results for reordered conjunctive query trees, and even worse, these results may be incomplete. Clearly, this represents a significant usability problem, as the explanations developers get may be partial and developers have to worry about the query tree representation of their query, losing the advantage of using a declarative query language. As remedy to this problem, in [5][17], we propose to capture query based answers of Why-Not questions through operator polynomial and we devised an algorithm called Ted that produces the same complete query-based explanations for reordered conjunctive query trees.

6.5. Social Data Management and Crowdsourcing

In [12], we focused on the issue of defining models and metrics for reciprocity in signed graphs. In unsigned directed networks, reciprocity quantifies the predisposition of network members in creating mutual connections. On the other hand, this concept has not yet been investigated in the case of signed graphs. We capitalize on the graph degeneracy concept to identify subgraphs of the signed network in which reciprocity is more likely to occur. This enables us to assess reciprocity at a global level, rather than at an exclusively local one as in existing approaches. The large scale experiments we perform on real world data sets of trust networks lead to both interesting and intuitive results. We believe these reciprocity measures can be used in various social applications such as trust management, community detection and evaluation of individual nodes. The global reciprocity we define in this paper is closely correlated to the clustering structure of the graph, more than the local reciprocity as it is indicated by the experimental evaluation we conducted.

As initial step towards better answering information needs in applications managing social content that is structured and possibly enriched with semantic annotations, in [19], we present a preliminary data model and an approach for answering queries over structured, social and semantic-rich content, taking into account all dimensions of the data in order to return the most meaningful results.

OPALE Project-Team

6. New Results

6.1. Highlights of the Year

Paola Goatin was awarded the *"Prix Inria - Académie des sciences du jeune chercheur"*.

6.2. Mathematical analysis and control of macroscopic traffic flow models

6.2.1. Vehicular traffic

Participants: Enrico Bertino, Guillaume Costeseque, Maria Laura Delle Monache, Paola Goatin, Sheila Scialanga, Alexandre Bayen [UC Berkeley, CA, USA], Sebastien Blandin [IBM Research Collaboratory, Singapore], Christophe Chalons [LJLL, UP7].

In collaboration with UC Berkeley, and as part of the Associated Team ORESTE activity (see http://www-sop.inria.fr/members/Paola.Goatin/ORESTE/index.html), we have considered the System Optimal Dynamic Traffic Assignment problem with Partial Control (SO-DTA-PC) for general road networks with horizontal queuing. The goal of which is to optimally control any subset of the networks agents to minimize the total congestion of all agents in the network. We adopt a flow dynamics model that is a Godunov discretization of the Lighthill-Williams-Richards (LWR) partial differential equation with a triangular flux function and a corresponding multi-commodity junction solver. Full Lagrangian paths are assumed to be known for the controllable agents, while we only assume knowledge of the aggregate turn ratios for the non-controllable (selfish) agents. We solve the resulting finite horizon non-linear optimal control problem using the discrete adjoint method, see [75].

As part of our TRAM3 activity and in collaboration with C. Chalons (UVSQ), we designed a new finite volume conservative algorithm to track the trajectory of a bus in the surrounding traffic using a locally non-uniform moving mesh, see [70].

In collaboration with S. Blandin (IBM), we proved the existence and stability of entropy weak solutions of a scalar conservation law with non-local flux arising in traffic flow modeling. The result is obtained providing accurate L^∞, BV and L^1 estimates for the sequence of approximate solutions constructed by an adapted Lax-Friedrichs scheme.

In collaboration with the University of Mannheim and in the framework of the PHC Procope project "Transport Networks Modeling and Analysis", we studied how to manage variable speed limits combined with coordinated ramp metering within the framework of the LWR network model. Following a "first discretize then optimize" approach, we derived the first order optimality system and explained the switch of speeds at certain fixed points in time and the boundary control for the ramp metering as well. Sequential quadratic programming methods are used to solve the control problem numerically. For application purposes, we present experimental setups where variable speed limits are used as a traffic guidance system to avoid traffic jams on highway interchanges and on-ramps, see [71].

Finally, E. Bertino internship was devoted to uncertainty quantification in macroscopic traffic flow models.

6.2.2. Crowd motion

Participants: Aekta Aggarwal, Régis Duvigneau, Paola Goatin, Matthias Mimault, Rinaldo M. Colombo [Brescia University, Italy].

A. Aggarwal postdoc is devoted to the analytical and numerical study of systems of conservation laws with non-local fluxes in several space dimensions. In collaboration with R.M. Colombo, we presented a Lax-Friedrichs type algorithm to numerically integrate this class of systems. The convergence of the approximate solutions was proved, also providing the existence of solution in a slightly more general setting than in other results in the current literature. An application to a crowd dynamics model is considered. This numerical algorithm is then used to test the conjecture that as the convolution kernels converge to a Dirac δ, the nonlocal problem converges to its non-nonlocal analogue.

M. Mimault is working on scalar conservation laws with non-local flow in two space dimensions. These equations are meant to model crowd motion, where the movement direction of each pedestrian depends on a weighted mean of the crowd density around him. In particular, M. Mimault is implementing a finite volume numerical scheme which will be used for flow optimization purposes.

The above researches were partially funded by the ERC Starting Grant "TRAM3 - Traffic management by macroscopic models".

6.3. Optimum design and control in fluid dynamics and its couplings

In computational sciences for physics and engineering, Computational Fluid Dynamics (CFD) are playing one of the major roles in the scientific community to foster innovative developments of numerical methodologies. Very naturally, our expertise in compressible CFD has led us to give our research on numerical strategies for optimum design a particular, but not exclusive focus on fluids.

The framework of our research aims to contribute to numerical strategies for PDE-constrained multi-objective optimization, with a particular emphasis on CPU-demanding computational applications in which the different criteria to be minimized (or reduced) originate from different physical disciplines that share the same set of design variables. These disciplines are often fluids, as a primary focus, coupled with some other disciplines, such as structural mechanics.

Our approach to *competitive optimization* is focused on the two-discipline problem. It is based on a particular construction of *Nash games*, relying on a *split of territory* in the assignment of individual strategies. A methodology has been proposed for the treatment of two-discipline optimization problems in which one discipline, the primary discipline, is preponderant, or fragile. Then, it is recommended to identify, in a first step, the optimum of this discipline alone using the whole set of design variables. Then, an orthogonal basis is constructed based on the evaluation at convergence of the Hessian matrix of the primary criterion and constraint gradients. This basis is used to split the working design space into two supplementary subspaces to be assigned, in a second step, to two virtual players in competition in an adapted Nash game, devised to reduce a secondary criterion while causing the least degradation to the first. The formulation has been proved to potentially provide a set of Nash equilibrium solutions originating from the original single-discipline optimum point by smooth continuation, thus introducing competition gradually [65] (see also subsectionsubsect:helico).

Our approach to *cooperative optimization*, in theory, is not limited in number of objective functions. It is based on a result of convex analysis established for a general unconstrained multiobjective problem in which all the gradients are assumed to be known. The theorem [66] states that in the convex hull of the gradients, there exists a unique vector of minimal norm, ω; if it is nonzero, the vector ω is a descent direction common to all criteria; otherwise, the current design point is Pareto-stationary. This result led us to generalize the classical steepest-descent algorithm by using the vector ω as search direction. We refer to the new algorithm as the multiple-gradient descent algorithm (MGDA). The MGDA yields to a Pareto-stationary point, and in practice actual Pareto-optimality is verified a posteriori.

The two approaches have been combined to explore the Pareto front segment-wise as illustrated on Figure 2 .

6.3.1. Multiple-Gradient Descent Algorithm (MGDA)

Participants: Jean-Antoine Désidéri, Régis Duvigneau, Camilla Fiorini, Matteo Giacomini, Abderrahmane Habbal, Adrien Zerbinati.

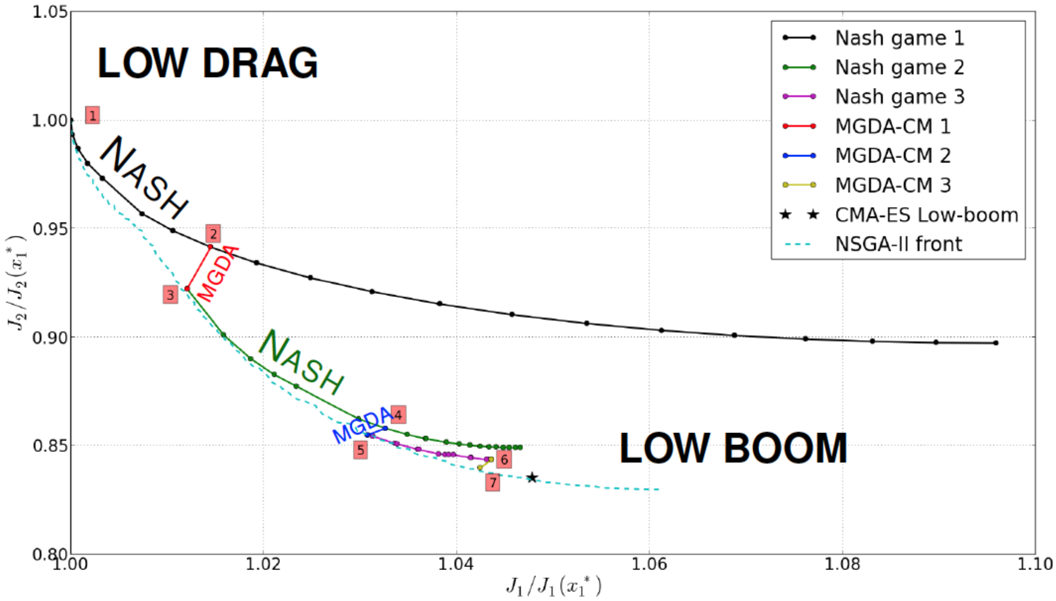


*Figure 2. **Two-discipline optimization of a generic geometry of a supersonic aircraft, for concurrent drag and sonic-boom reduction** (from A. Minelli's doctoral thesis, 2013). The wave drag is calculated by the ONERA elsA code in 3D finite-volume Eulerian flow mode over a 6M-node mesh and the sonic boom using a three-layer approach. The Nash-game paths have been devised by appropriate territory splitting in order to be tangent to the Pareto front, and they are interrupted whenever the Pareto-stationarity condition is judged excessively violated. The MGDA paths converge rapidly back to the front. The simulation demonstrates how the two algorithms complement each other and provide a potential for a piecewise description of the Pareto front, evaluated more economically than a stochastic algorithm operating on a large population.*

6.3.1.1. Theory and numerical experimentation of the MGDA construction

In multi-objective optimization, the knowledge of the Pareto set provides valuable information on the reachable optimal performance. A number of evolutionary strategies (PAES, NSGA-II, etc), have been proposed in the literature and proved to be successful to identify the Pareto set. However, these derivative-free algorithms are very demanding in terms of computational time. Today, in many areas of computational sciences, codes are developed that include the calculation of the gradient, cautiously validated and calibrated [66].

The notion of Pareto-stationarity, originally established to be a necessary condition of optimality in differentiable multi-objective optimization of unconstrained problems, has been extended to problems subject to equality constraints. On this basis, we were able to establish that by augmenting, in a classical manner, the objective-functions of a penalty term equal to the square of the constraint violation, and applying the standard MGDA to it, would result in converged solutions that are Pareto-stationary in the extended sense. Numerical experimentation on this is on-going.

6.3.1.2. Meta-model-assisted CFD optimization by MGDA

Using MGDA in a multi objective optimization problem requires the evaluation of a large number of points with regard to criteria, and their gradients. In the particular case of a CFD problems, each point evaluation is very costly since it involves a flow computation, possibly the solution of an adjoint-equation. To alleviate this difficulty, we have proposed to construct meta-models of the functionals of interest (lift, drag, etc) and to calculate approximate gradients by local finite differences. These meta-models are updated throughout the convergence process to the evaluation of the new design points by the high-fidelity model, here the 3D compressible Euler equations.

This variant of MGDA has been tested successfully over several aerodynamic shape optimization problems : lift concurrently with drag optimization for transonic aircraft; drag (under lift constraint) concurrently with sonic boom reduction for 3D supersonic configuration (at ONERA); drag (under lift constraint) concurrently with mass reduction for transport aircraft (at ONERA) [56].

6.3.1.3. Exact shape gradients

MGDA has successfully been tested over a two-objective optimization problem governed by two-dimensional elasticity. The deformation of a plate is calculated using an isogeometric approximation and compliance derived from it. The exact parametric shape gradient is calculated, yielding the gradient of the objective function in two antagonistic situations differing by the loading. Pareto-fronts are thus identified [68].

6.3.1.4. Optimization of an unsteady system using a multiobjective formulation

An approach has been developed to solve optimization problems in which the functional that has to be minimized is time dependent. In the literature, the most common approach when dealing with unsteady problems, is to consider a time-average criterion. However, this approach is limited since the dynamical nature of the state is neglected. Our alternative consists in building a set of cost functionals by evaluating a single criterion at different sampling times. In this way, the optimization of the unsteady system is formulated as a multi-objective optimization problem, solved using an appropriate descent algorithm (MGDA). Moreover, we also consider a hybrid approach in which the set of cost functionals is built by doing a time-average operation over multiple intervals. These strategies have been illustrated and applied to a non-linear unsteady system governed by a one-dimensional convection-diffusion-reaction partial differential equation [67].

6.3.1.5. Perspectives

MGDA offers the possibility to handle in a rational way several objective-functions for which gradients are known or approximated concurrently. This potential opens methodological paths to several themes of interest in high-fidelity simulation-based optimization: optimization of complex systems whose performance is evaluated w.r.t. several criteria originating from different, coupled disciplines; optimization under uncertainties, by introducing sensitivities as additional objectives; optimization of time-dependent systems, such as optimization of flow-control devices that generate a periodic flow (see next subsection), by converting the problem into a multipoint problem by time-discretization of the time and parameter-dependent functional (as above); etc.

6.3.2. Flow control

Participants: Régis Duvigneau, Jérémie Labroquère, Emmanuel Guilmineau [Ecole Centrale de Nantes].

Shape optimization methods are not efficient to improve the performance of fluid systems, when the flow is characterized by a strong unsteadiness related to a massive detachment. This is typically the case for the flow around an automotive body or a wing in stall condition. To overcome this difficulty, flow control strategies are developed, that aim at manipulating vortex dynamics by introducing some active actuators, such as periodic blowing/suction jets. In this context, the choice of the control parameters (location, amplitude, frequency) is critical and not straightforward. Therefore, we develop a methodology to determine optimal control parameters by coupling the simulation of unsteady actuated flows with optimization algorithms. Two research axes have been considered :

- the resolution of the unsteady sensitivity equations derived from the state equations, to exhibit the dependency of the flow dynamics with respect to the control and apply an unsteady gradient-based approach[67] ;
- the optimization of control parameters using a statistical metamodel-based strategy [39].

In this perspective, unsteady Reynolds Averaged Navier-Stokes equations are solved, with some turbulence closures. Different models for synthetic jet have been implemented to simulate the actuation, and then validated for different turbulence closures .

Specific developments have be carried out in the metamodel-based optimizer to include a noise term into Gaussian Process model, which is used to filter errors arising from unsteady simulations. A systematic assessment of modeling and numerical errors has been archived for a backward facing step test-case, with the objective of controlling the re-attachment point location[46], [58].

This activity is conducted in collaboration with the CFD team of Ecole Centrale de Nantes.

6.3.3. Adjoint-based mesh quality control

Participants: Jean-Antoine Desideri, Maxime Nguyen-Dinh [ONERA doctoral student], Jacques Peter [Research Engineer, ONERA/DSNA], Renaud Sauvage [Airbus France], Mathieu Meaux [EADS IW].

In his doctoral thesis [29], Nguyen Dinh has studied mesh adaptation methods based on the total derivatives of aerodynamic outputs with respect to mesh coordinates by the discrete adjoint method. Firstly, mesh adaptation methods have been devised for Eulerian flows. Zones to be refined are detected using a sensor based on the total derivative, and numerical experiments confirmed the adequacy of the approach. Secondly, the method was extended to the Reynolds-averaged Navier equations (RANS) and thirdly demonstrated for 3D industrial configurations [53].

6.3.4. Helicopter rotor blade optimization in both situations of hovering and forward flight

Participants: David Alfano [Airbus Helicopter], Michel Costes [Research Engineer, ONERA/DAAP], Jean-Antoine Désideri, Arnaud Le Pape [Research Engineer, ONERA/DAAP], Enric Roca Leon.

E. Roca Leon has conducted a CIFRE thesis at ONERA DAAP supported by Airbus Helicopter (Marignane) [34]. This thesis follows the doctoral thesis of A. Dumont in which the adjoint-equation approach was used to optimize a rotor blade in hovering flight. The goal of this new thesis is to solve a two-objective optimization problem in which the hovering-flight criterion is considered preponderant, but a new criterion that takes into account the forward-flight situation is also introduced, concurrently. The second criterion is the power necessary to maintain the forward motion. The first phase of thesis work has been devoted to the set up of a hierarchy of models from low to high fidelity, in order to calibrate appropriate functional criteria. Then, actual two-objective optimizations are conducted via our Nash game approach to competitive optimization with territory splitting based on reduced Hessian diagonalization. Successful optimization has been realized involving 16 geometrical parameters to reduce the power in forward motion while maintaining sub-optimality of the drag in hover [55] [64] [65].

6.4. Isogeometric analysis and design

Participants: Régis Duvigneau, Asma Gdhami, Bernard Mourrain [Galaad Project-Team], Bernd Simeon [Tech. Univ. of Kaiserslautern].

Design optimization stands at the crossroad of different scientific fields (and related software): Computer-Aided Design (CAD), Computational Fluid Dynamics (CFD) or Computational Structural Dynamics (CSM), parametric optimization. However, these different fields are usually not based on the same geometrical representations. CAD software relies on Splines or NURBS representations, CFD and CSM software uses grid-based geometric descriptions (structured or unstructured), optimization algorithms handle specific shape parameters. Therefore, in conventional approaches, several information transfers occur during the design phase, yielding approximations that can significantly deteriorate the overall efficiency of the design optimization procedure. Moreover, software coupling is often cumbersome in this context.

The isogeometric approach proposes to definitely overcome this difficulty by using CAD standards as a unique representation for all disciplines. The isogeometric analysis consists in developing methods that use NURBS representations for geometric modeling, computational domain description and solution basis functions. Using such a unique data structure allows to compute the solution on the exact geometry (not a discretized geometry), obtain a more accurate solution (high-order approximation), reduce spurious numerical sources of noise that deteriorate convergence, avoid data transfers between the software. Moreover, NURBS representations are naturally hierarchical and allows to define multi-level algorithms for solvers as well as optimizers.
In this context, a collaborative work has also been carried out with the Technical University of Kaiserslautern, concerning the computation of shape gradients for linear elasticity problems[42], [68]. Moreover, the doctoral thesis of Asma Gdhami, in collaboration with ENIT in Tunisia, has started and concerns the development of isogeometric schemes for hyperbolic systems.

6.5. Optimum design in structural mechanics

6.5.1. Shape Optimization in Multidisciplinary Non-Linear Mechanics

Participants: Aalae Benki, Jean-Antoine Désidéri, Abderrahmane Habbal, Gael Mathis [ArcelorMittal, CRAA].

In collaboration with the ArcelorMittal's Center for Research in Automotive and Applications (CRAA), we study the multidisciplinary shape and parameter design of highly non linear mechanical 2D and 3D structures. We have developed methods adapted to the approximation of Pareto Fronts such as Normal Boundary Intersection NBI and Normalized Normal Constraint Method NNCM. Due to the time consuming cost evaluation, the use of cheap to evaluate surrogate models is mandatory. We have studied the consistency of the approach NBI or NNCM plus surrogates, which turned out to be successful for a broad panel of standard mathematical benchmarks. The coupling is successfully applied to a small scale industrial case, namely the shape optimization of a can bottom vis à vis dome reversal pressure and dome growth criteria. We have then defined a Nash game between criteria where the latter are approximated by the RBF metamodels. First, we validated the computation of a Nash equilibrium for mathematical functions, then we computed Nash equilibria for the small scale industrial case of the shape optimization of the can bottom.

Then, we considered the 3D problem of an automotive twist beam. In this 3D case, we aim to Pareto-optimal shapes for two objectives, the first being to minimize the Von-Mises strain to guarantee the formability of the twist beam, and the second being to maximize the stiffness. For solution with higher stiffness than the initial one, we could decrease the thickness to obtain a mass reduction with the same end-user properties.

We also introduced, to our knowledge for the first time in the structural optimization area, the notion of Kalai-Smorodinky equilibria which is aimed at the selection of equilibria among Pareto-optimal solutions. We applied this notion of equilibria to both industrial cases, and compared the results to Nash equilibria.

6.5.2. Optimization of Addendum Surfaces in Stamping

Participants: Fatima Zahra Oujebbour, Rachid Ellaia, Abderrahmane Habbal, Ziheng Zhao.

Within the OASIS Consortium (ArcelorMittal, ErDF, Inria, UTC, EURODECISION, ESILV, NECS, Delta-CAD, SCILAB-DIGITEO), Opale Project leads the Optimization task. Our aim is to develop decentralized decision-making algorithms dedicated to find efficient solutions (Pareto optimal) in a complex multidisciplinary framework (forming, stamping, welding non-linear processes, spring-back, vibration, in-function linear processes, crash and fatigue non linear and non differentiable processes) for several (between three and five) criteria. An important difficulty when trying to identify the Pareto Front, even when using adapted methods such the Normal Boundary Intersection, is that the criteria involved (thanks to the high nonlinearity in the mechanical models) exhibit many local optima. So one must use global optimization methods. We have studied the hybrid approach Simulated Annealing with Simultaneous Perturbation SASP for a suite of mathematical test-cases. To envisage the application of our method to the complex CPU time consuming stamping process, we lead an intermediate phase dedicated to the validation of the SASP method for the minimization of the spring-back that follows the stamping of a metal sheet, the design variable being the process parameters (two then four parameters). Then, we considered the more complex shape design of the initial blank. The initial blank design is a critical step in stamping design procedure, therefore it should be optimally designed. Our aim is to find the optimal initial blank shape that avoids or at least minimizes the springback and failure flaws. For this study, the geometry of the blank contour is described by parametric spline curves. Seven control points (P1,...,P7) are used to define the spline curves in order to have a wide variety of geometries. The exact computational evaluation of our criteria, springback and failure, is very expensive (the FE model request around 45 min to predict these two criteria) and the design space is of quite high dimension. Therefore, we considered the recourse to the sparse grid interpolation. Optimization process based on sparse grid interpolation is an optimal alternative in which criteria can be approximated with a suitable interpolation formula that needs significantly less points than the full grid. the obtained metamodel using sparse grid interpolation needs less than 1s to predict springback and failure on the same computation machine. To find the optimal initial blank shape, it was decided to perform the optimization process using the obtained metamodel. The construction of the sparse grid interpolant was based on the Chebyshev Gauss-Lobatto grid type and using the polynomial basis functions. This technique achieves a good accuracy with a competitive number of grid points. The comparison of the obtained fronts shows that we can capture Pareto solutions by NBI and NNCM with fewer points than NSGAII which requires a large number of populations and several generations to obtain the Pareto front. [48] [49] [50]

6.6. Application of shape and topology design to biology and medicine

6.6.1. Assessing the ability of the 2D Fisher-KPP equation to model cell-sheet wound closure

Participants: Abderrahmane Habbal, Hélène Barelli [Univ. Nice Sophia Antipolis, CNRS, IPMC], Grégoire Malandain [Inria, EPI Morpheme].

We address in this joint collaboration the ability of the widely used Fisher-KPP equations to render some of the dynamical features of epithelial cell-sheets during wound closure.

Our approach is based on nonlinear parameter identification, in a two-dimensional setting, and using advanced 2D image processing of the video acquired sequences. As original contribution, we lead a detailed study of the profiles of the classically used cost functions, and we address the "wound constant speed" assumption, showing that it should be handled with care.

We study five MDCK cell monolayer assays in a reference, activated and inhibited migration conditions. Modulo the inherent variability of biological assays, we show that in the assay where migration is not exogeneously activated or inhibited, the wound velocity is constant. The Fisher-KPP equation is able to accurately predict, until the final closure of the wound, the evolution of the wound area, the mean velocity of the cell front, and the time at which the closure occurred. We also show that for activated as well as for inhibited migration assays, many of the cell-sheet dynamics cannot be well captured by the Fisher-KPP model. Original unexplored utilizations of the model such as wound assays classification based on the calibrated diffusion and proliferation rate parameters is ongoing. [47]

6.7. Distributed Systems

6.7.1. High-Performance manipulation and storage of e-Science data

Participants: Benoit Lange, Toan Nguyen.

The work carried in previous years on distributed High-Performance Computing for e-Science workflows has enlightened the need for appropriate tools and methods to manage petabyte and exabyte volumes of data. This has been the focus of the work carried by Benoit Lange during his Post-Doc position in 2014. It was dedicated to the definition and prototyping of a large-scale HPC platform to support the execution of application solvers, efficient storage and management of large-volumes of data produced by the simulation applications and the visualization of their results on high-end graphics workstations. This platform also includes analytics software to produce specific results corresponding to the user queries. This is based on the Hadoop ecosystem [59]. Is is central for the communication between the dedicated HPC nodes running the solvers and the visualization nodes interfacing the end-users. It includes high-speed storage with dedicated file systems on specific nodes, and long-term storage for reference data using magnetic juke-boxes that store petabytes of application data. This work is supported by an FP7 project in which Inria is responsible for the Data Management work-package (Call FP7-2013-ICT-11, Grant 619439, 2014-2016). The partners of the project, named VELaSSCo (Visualization for Extremely Large Scale Scientific Computing), are : CIMNE (SP, coordinator), JOTNE and SINTEF (No), ATOS (SP), Fraunhofer IGD (D) and the University of Edinburg (UK).

ORPAILLEUR Project-Team

6. New Results

6.1. Highlights of the Year

As highlights of the year, we would like to mention several elements, an award in a competition and a best paper. In addition we would like to also mention the importance gained by two other papers.

- Yen Low, a postdoctoral fellow from Stanford and Adrien Coulet (Orpailleur team) jointly developed a prototype named *Whypothesis?* whose goal is to provide explanations on drug side effects for which the molecular mechanism remains unknown. This prototype won the "Best Application Award" at the 2014 NCBO Hackathon (National Center for Biomedical Ontology), held at Stanford University, April 26-27 (http://www.bioontology.org/2014_NCBO_Hackathon).
- The paper [2] describing a first and original proposition for combining pattern structures and relational concept analysis won the best paper award at the International Conference on Formal Concept Analysis in Cluj-Napoca, Romania.
- The paper [10] published in Nucleic Acids Research describes the latest version of KBDOCK, which has had over 12,000 non-duplicate visitors since 2011.
- The paper [44] on polypharmacology represents a nice collaboration with Harmonic Pharma, and it was used for the cover issue of Journal Chemical Information (http://pubs.acs.org/toc/jcisd8/54/3).

BEST PAPER AWARD :

[56] **Formal Concept Analysis - 12th International Conference - Proceedings**. V. CODOCEDO, A. NAPOLI.

6.2. The Mining of Complex Data

Participants: Mehwish Alam, Aleksey Buzmakov, Melisachew Chekol, Victor Codocedo, Adrien Coulet, Elias Egho, Nicolas Jay, Florence Le Ber, Ioanna Lykourentzou, Luis-Felipe Melo, Amedeo Napoli, Chedy Raïssi, Mohsen Sayed, My Thao Tang, Yannick Toussaint.

Keywords: formal concept analysis, relational concept analysis, pattern structures, pattern mining, association rule, graph mining, sequence mining, biclustering

Formal Concept Analysis and pattern mining are suitable symbolic methods for KDDK, that may be used for real-sized applications. Global improvements are carried on the scope of applicability, the ease of use, the efficiency of the methods, and on the ability to fit evolving situations. Accordingly, the team is extending these symbolic data mining methods for working on complex data (e.g. textual documents, biological, chemical or medical data), involving objects with multi-valued attributes (e.g. domains or intervals), n-ary relations, sequences, trees and graphs.

6.2.1. FCA and Variations: RCA, Pattern Structures and Biclustering

There are a few extensions of FCA for handling contexts involving complex data formats, e.g. graphs or relational data. Among them, Relational Concept Analysis (RCA) is a process for analyzing objects described both by binary and relational attributes [2] [131]. The RCA process takes as input a collection of contexts and of inter-context relations, and yields a set of lattices, one per context, whose concepts are linked by relations. RCA can play has an important role in KDDK, especially in text mining [105].

Another extension of FCA is based on Pattern Structures (PS) [112], which allows to build a concept lattice from complex data, e.g. nominal, numerical, and interval data [119]. Since then, we worked on some experiments involving pattern structures, namely sequence mining [107], information retrieval and recommendation [58], [22], functional dependencies [50], [17] and biclustering [69], [41]. One of the next step is the adaptation of pattern structures to graph mining.

Moreover, the notion of similarity between objects is also closely related to pattern structures [102]: two objects are similar as soon as they share the same attributes (binary case) or attributes with similar values or the same description (at least in part). Combination of similarity and pattern structures is also under study, in particular for solving information retrieval and annotation problems.

In pattern mining as in FCA, one main problem is the volume of the output. One general idea is to extract patterns which show a "good behavior" w.r.t. a given measure. Such patterns or concepts are expected to have good characteristics and to provide effective knowledge. We have conducted in the framework of FCA a series of experiments on the so-called "stability measure", showing that this measure is able to detect significant patterns [54], [53].

Finally, there is also an on-going work relating FCA and semantic web. This work focuses on the classification within a concept lattice of the answers returned by SPARQL queries. The concept lattice is then used as an index for navigating and ranking the answers w.r.t. their content and interest for a given objective [47].

6.2.2. Sequence Mining

Sequence data is widely used in many applications. Consequently, mining sequential patterns and other types of knowledge from sequence data became an important data mining task. In the team, the main emphasis is on developing efficient mining algorithms for pattern classification problems. The most frequent sequences generally provide trivial information. When analyzing the set of frequent sequences with a low minimum support, the user is overwhelmed by millions of patterns.

In our recent work, we studied the notion of δ-freeness for sequences. While this notion has extensively been discussed for itemsets, our work is the first to extend it to sequences. We defined an efficient algorithm devoted to the extraction of δ-free sequential patterns. We presented the advantage of the δ-free sequences and highlighted their importance when building sequence classifiers, and we showed how they can be used to address the feature selection problem in statistical classifiers which optimizes both accuracy and earliness of predictions [68].

6.2.3. Mining and Understanding Healthcare Trajectories

With the increasing burden of chronic illnesses, administrative health care databases hold valuable information that could be used to monitor and assess the processes shaping the trajectory of care of chronic patients. In this context, temporal data mining methods are promising tools, though lacking flexibility in addressing the complex nature of medical events. In the thesis work of Elias Egho [15], new algorithms were designed to extract patient trajectory patterns with different levels of granularity by relying on external taxonomies [62], [34]. The algorithms rely on the general FCA framework to formalize the general notion of multidimensional healthcare trajectories. There was also another work focusing on the similarity measure among sequences. An efficient and original similarity measure was design for that purpose [8].

6.2.4. Video Game Analytics

The video game industry has grown enormously over the last twenty years, bringing new challenges to the artificial intelligence and data analysis communities. We tackled this year the problem of automatic discovery of strategies in real-time strategy games through pattern mining. Such patterns are the basic units for many tasks such as automated agent design, but also to build tools for the professionally played video games in the electronic sports scene. We presented a new formalism within a sequential pattern mining approach and a novel measure, the balance measure, telling how a strategy is likely to win [51]. We experimented our methodology on a real-time strategy game that is professionally played in the electronic sport community and laid plans on a future collaboration with the MIT Game Lab.

6.2.5. KDDK in Text Mining

Ontologies help software and human agents to communicate by providing shared and common domain knowledge, and by supporting various tasks, e.g. problem-solving and information retrieval. In practice, building an ontology depends on a number of "ontological resources" having different types: thesaurus, dictionaries, texts, databases, and ontologies themselves. We are currently working on the design of a methodology based on FCA and RCA for ontology engineering from heterogeneous ontological resources. This methodology is based on both FCA and RCA, and was previously successfully applied in domains such as astronomy and biology.

In the framework of the ANR Hybride project (see 8.2.1.2), an engineer is implementing a robust system based on these previous research results, for preparing the way to new research directions involving trees and graphs. Moreover, we led a first successful experiment on extracting drug-drug interactions applying "lazy pattern structure classification" to syntactic trees. In addition, in his thesis work, Mohsen Sayed focused on extracting relations between named entities using graph mining methods applied to dependency graphs [67]. We are currently investigating how this approach can be generalized, i.e. how to detect a relation between complex expressions which are not previously recognized as named entities.

The notion of "Jumping Emerging Patterns" (JEP) previously used in chemistry [101], was updated and adapted in the context of text mining within the ANR Termith project. The objective is to design a learning method for filtering candidate terms within a full text and to decide whether an occurrence should be tagged as a term, i.e. a positive example, or as a simple word, i.e. a negative example. The method extracts from a training set all JEPs which are considered as hypotheses. To reduce the number of JEPs and to retain only the more significant JEPs from a linguistic point of view, JEPs are weighted and a constraint solver is used to verify the maximal coverage of the positive examples. Results are currently under evaluation.

6.3. KDDK in Life Sciences

Participants: Adrien Coulet, Marie-Dominique Devignes, Bernard Maigret, Gabin Personeni, David Ritchie, Malika Smaïl-Tabbone.

The Life Sciences constitute a challenging domain for KDDK. Biological data are complex from many points of views, e.g. voluminous, high-dimensional and deeply inter-connected. Analyzing such data is a crucial issue in health care, environment and agronomy. Besides, many bio-ontologies are available and can be used to enhance the knowledge discovery process. Accordingly, the research work of the Orpailleur team in KDDK applied to Life Sciences is in concern with the use of bio-ontologies to improve KDDK, and as well information retrieval, access to "Linked Open Data" (LOD) and data integration.

6.3.1. Inductive Logic Programming for Mining Linked Open Data

Increasing amounts of biomedical data provided as LOD offer novel opportunities for knowledge discovery in biomedicine. We proposed and published an approach for selecting, integrating, and mining LOD with the goal of discovering genes responsible for a disease [11]. The selection step relies on a set of choices made by a domain expert to isolate relevant pieces of LOD. Because these pieces are potentially not linked, an integration step is required to connect unlinked pieces. The resulting graph is subsequently mined using Inductive Logic Programming (ILP) that presents two main advantages. First, the input format compliant with ILP (first order logic) is close to the format of LOD (RDF triples). Second, domain knowledge can be added to this input and used during the induction step. We have applied this approach to the characterization of genes responsible for intellectual disability. For this real-world use case, we could evaluate ILP results and assess the contribution of domain knowledge. Our ongoing efforts explore how the combination of rules coming from distinct theories can improve the prediction accuracy [70] [16].

6.3.2. Analysis of biomedical data annotated with ontologies

Annotating data with concepts of an ontology is a common practice in the biomedical domain. Resulting annotations define links between data and ontologies that are key for data exchange, data integration and data analysis. Since 2011, we collaborate with the National Center for Biomedical Ontologies (NCBO) to develop a large repository of annotations named the NCBO Resource Index [118]. This repository contains annotations of 36 biomedical databases annotated with concepts of more than 200 ontologies of the BioPortal (http://bioportal.bioontology.org/). In the preceding years, we compared the annotations of a database of biomedical publications (Medline) with two databases of scientific funding (Crisp and ResearchCrossroads) to profile disease research [122]. One main challenge remains to develop a knowledge discovery approach able to mine correlations between annotations based on BioPortal ontologies, i.e. is it possible to discover interesting knowledge units within these annotations?

Then, we proposed an adaptation of FCA techniques, namely pattern structures, to explore the annotations of biomedical databases [108]. We considered documents of biomedical databases annotated with sets of ontological concepts as objects in a pattern structure. Corresponding annotations have been classified according to several dimensions, where a dimension is related to a particular aspect of domain knowledge. The pattern structure formalism was applied to classify these annotations, allowing to discover correlations between annotations but also lacks of completion in the annotations that could be fixed afterward. This adaptation of pattern structures opens many perspectives in term of ontology reengineering and knowledge discovery.

6.4. Structural Systems Biology

Participants: Marie-Dominique Devignes, Bernard Maigret, David Ritchie, Malika Smaïl-Tabbone.

Keywords: bioinformatics, chemistry, docking, knowledge discovery, screening, systems biology

Structural systems biology aims to describe and analyze the many components and interactions within living cells in terms of their three-dimensional (3D) molecular structures. We are currently developing advanced computing techniques for molecular shape representation, protein-protein docking, protein-ligand docking, high-throughput virtual drug screening, and knowledge discovery in databases dedicated to protein-protein interactions.

6.4.1. The Hex Protein Docking Program

Our *Hex* protein docking software is being more widely used than ever before. The unique polar Fourier correlation approach used in *Hex* [129] allows the expensive FFT part of its calculations to be greatly accelerated on modern graphics processors (GPUs) [130]. *Hex* is freely available for download for academic users at http://hex.loria.fr. A public GPU-powered server has also been created (http://hexserver.loria.fr) [123]. In the last four years, the server has performed some 63,700 docking runs, and the program has had some 37,000 downloads. The latest version of the program has been used successfully to dock symmetric dimers (unpublished results) in the international “CAPRI” docking experiment [115]. A manuscript on performing polar Fourier docking using symmetry constraints is in preparation with the Nano-D team at Inria Grenoble.

6.4.2. KBDOCK: Protein Docking Using Knowledge-Based Approaches

In order to explore the possibilities of using structural knowledge of protein-protein interactions, Anisah Ghoorah recently developed the KBDOCK system as part of her doctoral thesis project [116]. KBDOCK is available at http://kbdock.loria.fr. KBDOCK combines coordinate data from the Protein Data Bank [106] with the Pfam protein domain family classification [111] in order to describe and analyze all known protein-protein interactions for which the 3D structures are available. We have demonstrated the utility of KBDOCK [114] for template-based docking using 73 complexes from the Protein Docking Benchmark [117]. We recently presented results obtained using KBDOCK at the CAPRI conference on protein docking in Utrecht [115]. In late 2013, we updated KBDOCK with the latest data from Pfam and the Protein Data Bank. In 2014, an article describing the new version of KBDOCK was published in the special Database Issue of Nucleic Acids Research [10]. Since the KBDOCK web site (http://kbdock.loria.fr) was created in 2011, it has had over 12,000 distinct visitors.

6.4.3. Kpax: A New Algorithm for Multiple Flexible Protein Structure Alignments

We recently developed a new protein structure alignment approach called Kpax [128]. The approach exploits the fact that each amino acid residue has a carbon atom with a highly predictable tetrahedral geometry. This allows the local environment of each residue to be transformed into a canonical orientation, thus allowing easy comparison between the canonical orientations of residues within pairs of proteins using a novel scoring function based on Gaussian overlaps. The overall approach is two or three orders of magnitude faster than most contemporary protein structure alignment algorithms, while still being almost as accurate as the state-of-the-art TM-Align approach [134]. Kpax is now used heavily by the KBDOCK web server [10] to find structural templates for docking which might be beyond the reach of sequence-based homology modeling approaches. The Kpax program is also available for download at http://kpax.loria.fr/.

In 2014, the Kpax algorithm has been extended to allow flexible alignment and superposition of protein backbones and to perform multiple structure alignments, in analogy with multiple protein sequence alignments. Our early results show that incorporating backbone flexibility leads to much higher quality multiple alignments than can be achieved with existing algorithms.

6.4.4. Polypharmacology: Developing New Uses for Old Drugs

In 2010, Violeta Pérez-Nueno joined the Orpailleur team thanks to a Marie Curie Intra-European Fellowship (IEF) award to develop new virtual screening algorithms (DOVSA). The aim of this project was to advance the state of the art in computational virtual drug screening by developing a novel consensus shape clustering approach based on spherical harmonic (SH) shape representations [126].

In 2012, Violeta joined Harmonic Pharma, a LORIA spin-out company for drug re-purposing, and we have since continued our collaborations to develop new algorithms for drug discovery and drug re-purposing. The observation that many existing drugs may be used to treat more than one disease is often referred to as "polypharmacology." Our latest work on predicting polypharmacology uses a Gaussian clustering approach to identify groups molecules with similar three-dimensional shapes. This work was published in the Journal of Chemical Information and Modeling [44]. An illustration from this article was used to provide the cover page for the March 2014 issue of the journal (http://pubs.acs.org/toc/jcisd8/54/3).

6.5. Around the Taaable research project

Participants: Valmi Dufour-Lussier, Emmanuelle Gaillard, Florence Le Ber, Jean Lieber, Amedeo Napoli, Emmanuel Nauer.

Keywords: knowledge representation, description logics, classification-based reasoning, case-based reasoning, belief revision, semantic web

The Taaable project was originally created as a challenger of the Computer Cooking Contest (ICCBR Conference) [4] (http://taaable.fr). A candidate to this contest is a system whose goal is to solve cooking problems.

Beyond its participation to the CCC challenges, the Taaable project aims at federating various research themes: case-based reasoning (CBR), information retrieval, knowledge acquisition and extraction, knowledge representation, minimal change theory, ontology engineering, semantic wikis, text-mining, etc. CBR performs adaptation of recipes w.r.t. user constraints. The reasoning process is based on a cooking domain ontology (especially hierarchies of classes) and adaptation rules. The knowledge base is encoded within a semantic wiki containing the recipes, the domain ontology and adaptation rules.

Minimal change theory and belief revision can be used as tools to support adaptation in CBR, i.e. the source case is modified to be consistent with the target problem using a revision operator. Belief revision was applied to Taaable to compute ingredient substitutions and to adjust the ingredient quantities [65] using engines included in the Revisor library (see § 5.4.5).

As acquiring knowledge from experts is costly, a new approach was proposed to allow a CBR system to use partially reliable, non expert, knowledge from the Web for reasoning. This approach is based on a meta-knowledge model to manage knowledge reliability. This model represents notions such as belief, trust, reputation and quality, as well as their relationships and rules to evaluate knowledge reliability. The reliability estimation is used to filter knowledge with high reliability as well as to rank the results produced by the CBR system. Performing CBR with knowledge resulting from an e-community is improved by taking into account the knowledge reliability [64].

Taaable won in 2014 the CCC originality challenge for all the open resources that the Taaable team developed during the last years for the CBR community: WikiTaaable, a semantic wiki containing cooking domain knowledge, Tuuurbine, a generic ontology guided CBR engine over RDFS (see § 5.4.3), and Revisor, an adaptation engine implementing various revision operators (see § 5.4.5).

6.6. Some Results in Graph Theory

Participants: Miguel Couceiro, Amedeo Napoli, Chedy Raïssi, Jean-Sébastien Sereni, Mario Valencia.

Keywords: graph theory, extremal graph theory, coloring, clustering

6.6.1. Structural and extremal graph theory

Regarding graph coloring, a conjecture of Gera, Okamoto, Rasmussen and Zhang on set coloring was solved. A *set coloring* of a graph $G = (V, E)$ is a function $c : V \to \{1, ..., k\}$ such that whenever u and v are adjacent vertexes, it holds that $\{c(x) : x \text{ neighbor of } u\} \neq \{c(x) : x \text{ neighbor of } v\}$. In other words, there must be at least one neighbor of u that has a color not assigned to a neighbor of v, or *vice-versa*. The smallest k such that G admits a set coloring is the *set coloring number* $\chi_s(G)$. We confirmed the conjecture by proving that $\chi_s(G) \geq \lceil \log_2 \chi(G) \rceil + 1$, where $\chi(G)$ is the (usual) chromatic number of G. This bound is tight.

Works have been started on a 12-year-old conjecture by Heckman and Thomas about the fractional chromatic number of graphs with no triangles and maximum degree at most 3. This conjecture is actually a natural generalization of a fact established by Staton in 1979. Heckman and Thomas posits that in every graph with no triangles, maximum degree at most 3 and arbitrary weights on the vertexes, there exists an independent set of weight at least $5/14$ times the total weight of the graph.

Regarding extremal graph theory, two results have been obtained. The first one deals with permutation snarks, while the second one reads as follows.

For every 3-coloring of the edges of the complete graph on n vertexes, there is a color c and a set X of 4-vertexes such that at least $2n/3$ vertexes are linked to a vertex in X by an edge of color c.

This theorem is motivated by a conjecture of Erdős, Faudree, Gould, Gyárfás, Rousseau and Schelp from 1989, which asserts that X can be of size 3 only. However, they were only able to prove that X can be of size 22. Recently, Rahil Baber and John Talbot managed to build upon our work in a very nice article: adding a new idea to our argument, they managed to confirm the conjecture.

6.6.2. Graph theory and other fields

Interactions of graph theory with other topics (theoretical computer science, number theory, group theory, sociology and chemistry) have been considered. Most of them are still in progress and some are published. For instance, regarding distributed computing, the purpose of our work was to question the global knowledge each node is assumed to start with in many distributed algorithms (both deterministic and randomized). More precisely, numerous sophisticated local algorithm were suggested in the literature for various fundamental problems. Noticeable examples are the MIS algorithms and the $(\Delta + 1)$-coloring algorithms. Unfortunately, most known local algorithms are *non-uniform*, that is, they assume that all nodes know good estimations of one or more global parameters of the network, e.g., the number of nodes n. Our work provides a rather general method for transforming a non-uniform local algorithm into a uniform one. Furthermore, the resulting algorithm enjoys the same asymptotic running time as the original non-uniform algorithm. Our method applies to a wide family of both deterministic and randomized algorithms. Specifically, it applies to almost all of the state of the art non-uniform algorithms regarding MIS and Maximal Matching, as well as to many results concerning the coloring problem.

6.6.3. Algorithmic Graph Theory and Clustering

Since September 2013, Mario Valencia has obtained a two years invitation (namely Inria "Délégation”) for working at Inria Nancy – Grand Est, in the Orpailleur team, on graph theoretical aspects and data clustering. This research work consists in studying the modular decomposition techniques on the threshold graphs issues of the clustering process. The principal studied problem is known as the *Cluster Deletion Problem*: given a graph with real non negative edge weights, partition the vertexes into clusters (in this case cliques) in order to minimize the total weight of edges out of the clusters. Two papers were submitted to journals in 2014. In [94], we discovered a one-to-one correspondence between potential solutions of the cluster deletion problem and the minimum sum coloring problem, and use it to obtain a polynomial time algorithm to solve the cluster deletion problem in a special family of graphs called P_4-reducible graphs.

In [95], we studied the complexity of the cluster deletion problem on subclasses of chordal graphs and cographs. In particular, it is shown that the cluster deletion problem is NP-hard for unweighted chordal graphs and weighted cographs. Some polynomial-time solvable cases are also identified.

Moreover, the paper "b-coloring is NP-hard on co-bipartite graphs and polytime solvable on tree-cographs", has been accepted for publication in the journal *Algorithmica* [1].

6.6.4. Structural and Algebraic Graph Theory

We have also worked on the following topics. Golumbic, Lipshteyn and Stern proved that every graph can be represented as the edge intersection graph of paths on a grid, i.e., one can associate to each vertex of the graph a nontrivial path on a grid such that two vertexes are adjacent if and only if the corresponding paths share at least one edge of the grid. For a non-negative integer k, B_k-EPG graphs are defined as graphs admitting a model in which each path has at most k bends. Circular-arc graphs are intersection graphs of open arcs of a circle. It is easy to see that every circular-arc graph is B_4-EPG, by embedding the circle into a rectangle of the grid. We proved also that every circular-arc graph is B_3-EPG (paper submitted).

We have studied the k-tuple chromatic number of the Cartesian product of two graphs G and H in [96]. We have shown that there exist graphs G and H such that $\chi_k(G\Box H) > \max\{\chi_k(G), \chi_k(H)\}$ for $k \geq 2$. Moreover, we have also shown that there exist graph families such that, for any $k \geq 1$, the k-tuple chromatic number of their Cartesian product is equal to the maximum k-tuple chromatic number of its factors.

PANAMA Project-Team

6. New Results

6.1. Highlights of the Year

The EUSIPCO 2014 Best Student Paper Award was awarded to our joint paper [32] on dynamic screening for sparse regularization. .

A review paper on audio source separation, rooted in METISS/PANAMA know-how and contributions to this topic over the years, was published in the IEEE Signal Processing Magazine [25].

A new version of the Flexible Audio Source Separation Toolbox (FASST) was released in January 2014 and downloaded 300 times.
BEST PAPERS AWARDS :
[32] **European Signal Processing Conference EUSIPCO 2014**. A. BONNEFOY, V. EMIYA, L. RALAIVOLA, R. GRIBONVAL.

6.2. Recent results on sparse representations

Sparse approximation, high dimension, scalable algorithms, dictionary design, sample complexity

The team has had a substantial activity ranging from theoretical results to algorithmic design and software contributions in the field of sparse representations, which is at the core of the ERC project PLEASE (projections, Learning and Sparsity for Efficient Data Processing, see Section 8.2.1.1).

6.2.1. A new framework for sparse representations: analysis sparse models

Participants: Rémi Gribonval, Nancy Bertin, Srdan Kitic, Cagdas Bilen, Laurent Albera.

In the past decade there has been a great interest in a synthesis-based model for signals, based on sparse and redundant representations. Such a model assumes that the signal of interest can be composed as a linear combination of *few* columns from a given matrix (the dictionary). An alternative *analysis-based* model can be envisioned, where an analysis operator multiplies the signal, leading to a *cosparse* outcome. Within the SMALL FET-Open project, we initiated a research programme dedicated to this analysis model, in the context of a generic missing data problem (e.g., compressed sensing, inpainting, source separation, etc.). We obtained a uniqueness result for the solution of this problem, based on properties of the analysis operator and the measurement matrix. We also considered a number of pursuit algorithms for solving the missing data problem, including an ℓ^1-based and a new greedy method called GAP (Greedy Analysis Pursuit). Our simulations demonstrated the appeal of the analysis model, and the success of the pursuit techniques presented.

These results have been published in conferences and in a journal paper [100]. Other algorithms based on iterative cosparse projections [83] as well as extensions of GAP to deal with noise and structure in the cosparse representation have been developed, with applications to toy MRI reconstruction problems and acoustic source localization and reconstruction from few measurements [101].

Successful applications of the cosparse approach to sound source localization, audio declipping and brain imaging have been developed this year. In particular, we compared the performance of several cosparse recovery algorithms in the context of sound source localization [97] and showed its efficiency in situations where usual methods fail ([37], see paragraph 6.6.3). It was also shown to be applicable to the hard declipping problem [49]. Application to EEG brain imaging was also investigated and a paper was published in MLSP14 [28] (see paragraph 6.6.4).

6.2.2. Theoretical results on sparse representations

Participants: Rémi Gribonval, Anthony Bourrier, Pierre Machart, Yann Traonmilin, Gilles Puy.

Main collaboration: Mike Davies (University of Edinburgh), Patrick Perez (Technicolor R&I France), Tomer Peleg (The Technion)

Fundamental performance limits for ideal decoders in high-dimensional linear inverse problems: The primary challenge in linear inverse problems is to design stable and robust "decoders" to reconstruct high-dimensional vectors from a low-dimensional observation through a linear operator. Sparsity, low-rank, and related assumptions are typically exploited to design decoders which performance is then bounded based on some measure of deviation from the idealized model, typically using a norm. We characterized the fundamental performance limits that can be expected from an ideal decoder given a general model, ie, a general subset of "simple" vectors of interest. First, we extended the so-called notion of instance optimality of a decoder to settings where one only wishes to reconstruct some part of the original high dimensional vector from a low-dimensional observation. This covers practical settings such as medical imaging of a region of interest, or audio source separation when one is only interested in estimating the contribution of a specific instrument to a musical recording. We defined instance optimality relatively to a model much beyond the traditional framework of sparse recovery, and characterized the existence of an instance optimal decoder in terms of joint properties of the model and the considered linear operator [106], [105]. This year, noiseless and noise-robust settings were both considered in the journal paper [16]. We showed somewhat surprisingly that the existence of noise-aware instance optimal decoders for all noise levels implies the existence of a noise-blind decoder. A consequence of our results is that for models that are rich enough to contain an orthonormal basis, the existence of an L2/L2 instance optimal decoder is only possible when the linear operator is not substantially dimension-reducing. This covers well-known cases (sparse vectors, low-rank matrices) as well as a number of seemingly new situations (structured sparsity and sparse inverse covariance matrices for instance). We exhibit an operator-dependent norm which, under a model-specific generalization of the Restricted Isometry Property (RIP), always yields a feasible instance optimality and implies instance optimality with certain familiar atomic norms such as the ℓ^1 norm. Current work explores the existence of convex decoders for general union of subspaces models under generalized RIP assumptions, as well as conditions ensuring that random low-dimensional projections ensure the RIP even when the projection is from an infinite-dimensional space to a finite dimensional one. Envisioned applications are in compressive learning (see Section 6.4).

Connections between sparse approximation and Bayesian estimation: Penalized least squares regression is often used for signal denoising and inverse problems, and is commonly interpreted in a Bayesian framework as a Maximum A Posteriori (MAP) estimator, the penalty function being the negative logarithm of the prior. For example, the widely used quadratic program (with an ℓ^1 penalty) associated to the LASSO / Basis Pursuit Denoising is very often considered as MAP estimation under a Laplacian prior in the context of additive white Gaussian noise (AWGN) reduction.

In 2011 we obtained a result [85] highlighting the fact that, while this is *one* possible Bayesian interpretation, there can be other equally acceptable Bayesian interpretations. Therefore, solving a penalized least squares regression problem with penalty $\phi(x)$ need not be interpreted as assuming a prior $C \cdot \exp(-\phi(x))$ and using the MAP estimator. In particular, we showed that for *any* prior P_X, the minimum mean square error (MMSE) estimator is the solution of a penalized least square problem with some penalty $\phi(x)$, which can be interpreted as the MAP estimator with the prior $C \cdot \exp(-\phi(x))$. Vice-versa, for *certain* penalties $\phi(x)$, the solution of the penalized least squares problem is indeed the MMSE estimator, with a certain prior P_X. In general $dP_X(x) \neq C \cdot \exp\left(-\phi(x)\right)dx$. In 2013, we extended this result to general inverse problems [88], [86], [87]. This year, we worked on the characterization of such relations beyond the Gaussian noise model, with the objective of understanding whether similar results hold when the quadratic data-fidelity term is replaced with other convex losses.

6.2.3. Algorithmic and theoretical results on dictionary learning

Participants: Rémi Gribonval, Nancy Bertin, Srdan Kitic, Cagdas Bilen, Luc Le Magoarou, Melanie Ducoffe.

Main collaboration (theory for dictionary learning): Rodolphe Jenatton, Francis Bach (Equipe-projet SIERRA (Inria, Paris)), Martin Kleinsteuber, Matthias Seibert (TU-Munich), *Main collaboration (dictionary learning for gesture recognition): Anatole Lecuyer, Ferran Argelaguet (EPI HYBRID, Rennes)*

Theoretical guarantees for dictionary learning : An important practical problem in sparse modeling is to choose the adequate dictionary to model a class of signals or images of interest. While diverse heuristic techniques have been proposed in the litterature to learn a dictionary from a collection of training samples, there are little existing results which provide an adequate mathematical understanding of the behaviour of these techniques and their ability to recover an ideal dictionary from which the training samples may have been generated.

Beyond our pioneering work [89], [109] [6] on this topic, which concentrated on the noiseless case for non-overcomplete dictionaries, this year we obtained new results showing the relevance of an ℓ^1 penalized cost function for the locally stable identification of overcomplete incoherent dictionaries, in the presence of noise and outliers [54]. Moreover, we established new sample complexity bounds of dictionary learning and other related matrix factorization schemes (including PCA, NMF, structured sparsity ...) [55], [46], [38].

Learning computationally efficient dictionaries Classical dictionary learning is limited to small-scale problems. Inspired by usual fast transforms, we proposed a general dictionary structure that allows cheaper manipulation, and an algorithm to learn such dictionaries –and their fast implementation [50]. A preprint is available [56], a paper will appear at ICASSP 2015, and a journal paper is in preparation.

Operator learning for cosparse representations : Besides standard dictionary learning, we also considered learning in the context of the cosparse model. The overall problem is to learn a low-dimensional signal model from a collection of training samples. The mainstream approach is to learn an overcomplete dictionary to provide good approximations of the training samples using sparse synthesis coefficients. This famous sparse model has a less well known counterpart, in analysis form, called the cosparse analysis model. In this new model, signals are characterized by their parsimony in a transformed domain using an overcomplete analysis operator.

In specific situations, when prior information is available on the operator, it is possible to express it in parametric form and learn this parameter. For instance, in the sound source localization problem, we showed that the unknown speed of sound can be learned jointly in the process of cosparse recovery, under mild conditions. This work was presented at iTwist'14 workshop [48].

Dictionary learning for gesture modeling In collaboration with the HYBRID project-team (internship of Melanie Ducoffe), we explored the potential of dictionary learning in the context of motion tracking. Motion tracking technology, especially for commodity hardware, requires robust gesture recognition algorithms to fully exploit the benefits of natural user interfaces. We proposed a gesture recognition algorithm based on the sparse representation of motion data, with a learning phase consisting in learning a dictionary of basic gestures. A paper is in preparation.

6.3. Activities on waveform design for telecommunications

Participant: Rémi Gribonval.

Main collaboration: Marwa Chafii, Jacques Palicot, Carlos Bader (Equipe SCEE, Supelec, Rennes)

Peak to Average Power Ratio (PAPR), Orthogonal Frequency Division Multiplexing (OFDM), Generalized Waveforms for Multi Carrier (GWMC)

In the context of the TEPN (Towards Energy Proportional Networks) Comin Labs project (see Section 8.1.1.2), in collaboration with the SCEE team at Supelec (thesis of Marwa Chafii co-supervised by R. Gribonval), we investigated a problem related to dictionary design: the characterization of waveforms with low Peak to Average Power Ratio (PAPR) for wireless communications. This is motivated by the importance of a low PAPR for energy-efficient transmission systems. A first stage of the work consisted in characterizing the statistical distribution of the PAPR for a general family of multi-carrier systems, leading to a journal paper [17] and several conference communications [27], [33]. The work now concentrates on characterizing waveforms with optimum PAPR.

6.4. Emerging activities on compressive sensing, learning and inverse problems

Compressive sensing, compressive learning, acoustic wavefields, audio inpainting,

6.4.1. Audio inpainting

Participants: Rémi Gribonval, Nancy Bertin, Corentin Guichaoua, Srdan Kitic, Anh Tho Le.

Inpainting is a particular kind of inverse problems that has been extensively addressed in the recent years in the field of image processing. It consists in reconstructing a set of missing pixels in an image based on the observation of the remaining pixels. Sparse representations have proved to be particularly appropriate to address this problem. However, inpainting audio data has never been defined as such so far. A series of works about audio inpainting was initiated by the METISS team in the framework of the EU Framework 7 FET-Open project FP7-ICT-225913-SMALL (Sparse Models, Algorithms and Learning for Large-Scale data).

Building upon our previous contributions (definition of the audio inpainting problem as a general framework for many audio processing tasks, application to the audio declipping or desaturation problem, formulation as a sparse recovery problem [60]), new results were obtained this year to address the case of audio declipping with the competitive cosparse approach. Its promising results, especially when the clipping level is low, were confirmed experimentally by the formulation and use of a new algorithm named Cosparse Iterative Hard Tresholding, which is a counterpart of the sparse Consistent Iterative Hard Thresholding. These results were presented during the iTwist'14 workshop [49]. Additional experiments were performed (internship of Anh Tho Le) to confirm the results on a larger database and investigate optimal parameters (nature and redundancy of the dictionary, relaxation parameter for the cosparsity level).

Current and future works deal with developing advanced (co)sparse decomposition for audio inpainting, including several forms of structured sparsity (*e.g.* temporal and multichannel joint-sparsity), dictionary learning for inpainting, and several applicative scenarios (declipping, time-frequency inpainting).

6.4.2. Blind Calibration of Compressive Sensing systems

Participants: Rémi Gribonval, Cagdas Bilen, Gilles Puy.

Main collaborations: Gilles Chardon, Laurent Daudet (Institut Langevin)

We consider the problem of calibrating a compressed sensing measurement system under the assumption that the decalibration consists of unknown gains on each measure. We focus on blind calibration, using measures performed on a few unknown (but sparse) signals. A naive formulation of this blind calibration problem, using ℓ^1 minimization, is reminiscent of blind source separation and dictionary learning, which are known to be highly non-convex and riddled with local minima. In the considered context, when the gains are real valued and non-negative, we showed that in fact this formulation can be exactly expressed as a convex optimization problem, and can be solved using off-the-shelf algorithms. Numerical simulations demonstrated the effectiveness of the approach even for highly uncalibrated measures, when a sufficient number of (unknown, but sparse) calibrating signals is provided. We observed that the success/failure of the approach seems to obey sharp phase transitions [84]. We extended the framework to phase-only decalibration, using techniques revolving around low-rank matrix recovery [66], [65], [110], [64], and to joint phase and gain decalibration [15].

6.4.3. Compressive Gaussian Mixture estimation

Participants: Rémi Gribonval, Anthony Bourrier, Nicolas Keriven.

Main collaborations: Patrick Perez (Technicolor R&I France)

When fitting a probability model to voluminous data, memory and computational time can become prohibitive. In this paper, we propose a framework aimed at fitting a mixture of isotropic Gaussians to data vectors by computing a low-dimensional sketch of the data. The sketch represents empirical moments of the underlying probability distribution. Deriving a reconstruction algorithm by analogy with compressive sensing, we experimentally show that it is possible to precisely estimate the mixture parameters provided that the sketch is large enough. Our algorithm provides good reconstruction and scales to higher dimensions than previous probability mixture estimation algorithms, while consuming less memory in the case of numerous data. It also provides a privacy-preserving data analysis tool, since the sketch does not disclose information about individual datum it is based on [71], [69], [70]. This year, extensions to non-isotropic Gaussians, with new algorithms and preliminary applications to speaker verification have been conducted.

6.5. Recent results on tensor decompositions

Multi-linear algebra is defined as the algebra of q-way arrays ($q > 2$), that is, the arrays whose elements are addressed by more than two indices. The first works back as far as Jordan who was interested in simultaneously diagonalizing two matrices at a time [96]. It is noteworthy that such two matrices can be interpreted as both slices of a three-way array and their joint diagonalization can be viewed as Hitchcock's polyadic decomposition [92] of the associated three-way array. Other works followed discussing rank problems related to multi-way structures and properties of multi-way arrays. However, these exercices in multilinear algebra were not linked to real data analysis but stayed within the realm of mathematics. Studying three-way data really started with Tucker's seminal work, which gave birth to the three-mode factor analysis [114]. His model is now often referred to as the Tucker3 model. At the same moment, other authors focused on a particular case of the Tucker3 model, calling it PARAFAC for PARAllel FACtor analysis [91], and on the means to achieve such a decomposition, which will become the famous canonical decomposition [74]. In honor to Hitchcock's pionneer work, we will call it the Canonical Polyadic (CP) decomposition.

Achieving a CP decomposition has been seen first as a mere non-linear least squares problem, with a simple objective criterion. In fact, the objective is a polynomial function of many variables, where some separate. One could think that this kind of objective is easy because smooth, and even infinitely differentiable. But it turns out that things are much more complicated than they may appear to be at first glance. Nevertheless, the Alternating Least Squares (ALS) algorithm has been mostly utilized to address this minimization problem, because of its programming simplicity. This should not hide the inherently complicated theory that lies behind the optimization problem. Moreover, in most of the applications, actual tensors may not exactly satisfy the expected model, so that the problem is eventually an approximation rather than an exact decomposition. This may results in a slow convergence (or lack of convergence) of iterative algorithms such as the ALS one [98]. Consequently, a new class of efficient algorithms able to take into account the properties of tensors to be decomposed is needed.

6.5.1. CP decomposition of semi-symmetric semi-nonnegative three-way arrays

Participant: Laurent Albera.

Main collaboration (Line search and trust region strategies): Julie Coloigner (LTSI, France), Amar Kachenoura (LTSI, France), Lotfi Senhadji (LTSI, France)

Main collaborations (Jacobi-like approaches): Lu Wang (LTSI, France), Amar Kachenoura (LTSI, France), Lotfi Senhadji (LTSI, France), Huazhong Shu (LIST, China)

We proposed new algorithms for the CP decomposition of semi-nonnegative semi-symmetric three-way tensors. In fact, it consists in fitting the CP model for which two of the three loading matrices are nonnegative and equal. Note that such a problem can also be interpreted as a nonnegative Joint Diagonalization by Congruence (JDC) problem.

Line search and trust region strategies

We first circumvented the nonnegativity constraint by means of changes of variable into squares, leading to a (polynomial) unconstrained optimization problem. Two optimization strategies, namely line search and trust region, were then studied. Regarding the former, a global plane search scheme was considered. It consists in computing, for a given direction, one or two optimal stepsizes, depending on whether the same stepsize is used in various updating rules. Moreover, we provided a compact matrix form for the derivatives of the objective function. This allows for a direct implementation of several iterative algorithms such as Conjugate Gradient (CG), Levenberg-Marquardt (LM) and Newton-like methods, in matrix programming environments like MATLAB. Note that the computational complexity issue was taken into account in the design phase of the algorithms, and was evaluated for each algorithm, allowing to fairly compare their performance.

Thus, various scenarios have been considered, aiming at testing the influence of i) an additive noise, which can stand for modeling errors, ii) the collinearity between factors, iii) the array rank and iv) the data size. The comparisons between our CG-like, Newton-like and LM-like methods (where semi-nonnegativity and semi-symmetry constraints are exploited), and classical CP algorithms (where no constraints are considered), showed that a better CP decomposition is obtained when these a priori are exploited, especially in the context of high dimensions and high collinearity. Finally, based on our numerical analysis, the algorithms that seem to yield the best tradeoff between accuracy and complexity are our CG_{2steps}-like and LM-like algorithms. This work was published in the Elsevier Linear Algebra and Applications journal [19].

Next, we considered an exponential change of variable leading to a different (non-polynomial) unconstrained optimization problem. Then we proposed novel algorithms based on line search strategy with an analytic global plane search procedure requiring new matrix derivations. Their performance was evaluated in terms of estimation accuracy and computational complexity. The classical ELS-ALS [108] and LM [112] algorithms without symmetry and nonnegativity constraints, and the ACDC algorithm [115] where only the semi-symmetry constraint is imposed, were tested as reference methods. Furthermore, the performance was also compared with our algorithms based on a square change of variable. The comparison studies showed that, among these approaches, the best accuracy/complexity trade off was achieved when an exponential change of variable was used through our ELS-ALS-like algorithm. This work was published in the Elsevier Signal Processing journal [18].

Jacobi-like approaches

The line search (despite the use of global plane search procedures) and trust region strategies may be sensitive to initialization, and generally require a multi-initialization procedure. In order to circumvent this drawback, we considered in this work Jacobi-like approaches, which are known to be less sensitive to initialization. Note that our line search and trust region approaches can then be used to refine the solution obtained by the latter.

More particularly, we formulated the high-dimensional optimization problem into several sequential polynomial and rational subproblems using i) a square change of variables to impose nonnegativity and ii) LU or QR matrix factorization for parameterization. The two equal nonnegative loading matrices are actually written as the Hadamard product of two equal matrices which can be factorized as the product of elementary matrices, each one depending on only one parameter.

The proposed approach reduces the optimization problem to the computation of the two equal nonnegative loading matrices only. The third loading matrix is algebraically derived from the latter. This requires an appropriate parameterization of the set of matrices whose inverse is nonnegative. This work was published in a journal paper [26]. Numerical experiments on simulated matrices emphasize the advantages of the proposed algorithms over classical CP and JDC techniques, especially in the case of degeneracies.

6.6. Source separation and localization

Source separation, sparse representations, tensor decompositions, semi-nonnegative independent component analysis, probabilistic model, source localization

6.6.1. A general framework for audio source separation

Participants: Frédéric Bimbot, Rémi Gribonval, Nancy Bertin.

Main collaboration: E. Vincent, Y. Salaün (EPI PAROLE, Inria Nancy); A. Ozerov, N.Q.K. Duong (Technicolor R&I France)

Source separation is the task of retrieving the source signals underlying a multichannel mixture signal.

About a decade ago, state-of-the-art approaches consisted of representing the signals in the time-frequency domain and estimating the source coefficients by sparse decomposition in that basis. These approaches rely only on spatial cues, which are often not sufficient to discriminate the sources unambiguously. Over the last years, we proposed a general probabilistic framework for the joint exploitation of spatial and spectral cues [102], which generalizes a number of existing techniques including our former study on spectral GMMs [61]. We showed how it could be used to quickly design new models adapted to the data at hand and estimate its parameters via the EM algorithm., and it became the basis of a large number of works in the field, including our own. In the last years, improvements were obtained through the use of prior knowledge about the source spatial covariance matrices [81], [95], [94], knowledge on the source positions and room characteristics [82], or a better initialization of parameters thanks to specific source localization techniques [68].

This accumulated progress lead to two main achievements this year, which show the maturity of our work and which will leverage its impact. First, a new version of the Flexible Audio Source Separation Toolbox, fully reimplemented, was released. It will provide the community with an efficient and ergonomic software, making available the tools from past years' research [58]. Second, we published an overview paper on recent and going research along the path of *guided* separation, *i.e.*, techniques and models allowing to incorporate knowledge in the process towards efficient and robust solutions to the audio source separation problem, in a special issue of IEEE Signal Processing Magazine devoted to source separation and its applications [25].

6.6.2. Towards real-world separation and remixing applications

Participants: Nancy Bertin, Frédéric Bimbot, Jules Espiau de Lamaestre, Anaik Olivero, Jérémy Paret, Nathan Souviraà -Labastie.

Emmanuel Vincent (EPI PAROLE, Inria Nancy)

While some challenges remain, work from previous years and our review paper on guided source separation [25] highlighted that progress has been made and that audio source separation is closer than ever to successful industrial applications, especially when some knowledge can be incorporated. This is exemplified by the contract with MAIA Studio, which reaches its end in December 2014 and showed in particular how user input or side information could raise source separation tools to efficient solutions in real-world applications.

In this context, new tools were developed this year. The introduction of manually-tuned parameters in the automated separation process, which modifies the Wiener filtering coefficients obtained from estimation of the mixtures covariance matrices, allows to find a better trade-off between artifacts and interferences. In order to ensure high audio quality for such applications, some user-guided corrections remain necessary even after an automatic pre-separation ; to this end, we developed an improved display (based on cepstrum and automatic constrast adaptation) and semi-automatic selection and suppression tools in the time-frequency domain. Those tools take as few inputs as possible from the user and their result can be ergonomically adjusted from the baseline output to a manually fine-tuned area, in a very small operating time. We also proposed tools to suppress a time-frequency area and replace it by content extracted from its context, reducing the perceptual impact of the suppression.

In some applicative contexts of source separation, several mixtures are available which contain similar instances of a given source. We have designed a general framework for audio source separation guided by multiple audio references, where each audio reference is a mixture which is supposed to contain at least one source similar to one of the target sources. Deformations between the sources of interest and their references are modeled in a general manner. A nonnegative *matrix co-factorization* algorithm is used which allows sharing of information between the considered mixtures. We have experimented our algorithm on music plus voice mixtures with music and/or voice references. Applied on movies and TV series data, the algorithm improves the signal-to-distortion ratio (SDR) of the sources of lowest intensity by 9 to 12 decibels with respect to original mixture [40]

6.6.3. Acoustic source localization

Participant: Nancy Bertin, Srdan Kitic, Laurent Albera, Nancy Bertin, Rémi Gribonval.

Main collaborations (audio-based control for robotics): Aly Magassouba and François Chaumette (Inria, EPI LAGADIC, France)

Acoustic source localization is, in general, the problem of determining the spatial coordinates of one or several sound sources based on microphone recordings. This problem arises in many different fields (speech and sound enhancement, speech recognition, acoustic tomography, robotics, aeroacoustics...) and its resolution, beyond an interest in itself, can also be the key preamble to efficient source separation. Common techniques, including beamforming, only provides the *direction of arrival* of the sound, estimated from the *Time Difference of Arrival* (*TDOA*) [68]. This year, we have particularly investigated alternative approaches, either where the explicit localization is not needed (audio-based control of a robot) or, on the contrary, where the exact location of the source is needed and/or TDOA is irrelevant (cosparse modeling of the acoustic field).

Implicit localization through audio-based control for robotics

In robotics, the use of aural perception has received recently a growing interest but still remains marginal in comparison to vision. Yet audio sensing is a valid alternative or complement to vision in robotics, for instance in homing tasks. Most existing works are based on the relative localization of a defined system with respect to a sound source, and the control scheme is generally designed separately from the localization system.

In contrast, the approach that we started investigating this year focuses on a sensor-based control approach. We proposed a new line of work, by considering the hearing sense as a direct and real-time input of closed loop control scheme for a robotic task. Thus, and unlike most previous works, this approach does not necessitate any explicit source localization: instead of solving the localization problem, we focus on developing an innovative modeling based on sound features. To address this objective, we placed ourselves in the sensor-based control framework, especially visual servoing (VS) that has been widely studied in the past [76].

From now on, we have established an analytical model linking sound features and control input of the robot, defined and analyzed robotic homing tasks involving multiple sound sources, and validated the proposed approach by simulations. This work is mainly lead by Aly Magassouba, whose Ph.D. is co-supervised by Nancy Bertin and François Chaumette. A conference paper presenting these first results was submitted to ICRA 2015. Future work will include real-world experiments with the robot Romeo from Aldebaran Robotics.

Cosparse modeling of the acoustic field

Cosparse modeling is particularly attractive when the signals of interest satisfy certain physical laws that naturally drive the choice of an analysis operator, which is the case for the acoustic field, ruled by the wave equation. Unlike usual localization techniques such as beamforming or TDOA-based direction estimation, which generally consider reverberation as an adverse condition, the cosparse modeling of sound propagation has also the interest of considering reverberation as a source of additional information for the localization task. Eventually, it can provide a full coordinate localization of the sources, and not only their direction of arrival.

Building upon our previous results on cosparse modeling and recovery algorithms for the wave equation [97], we have obtained additional evidence of the interest of this approach. In particular, we have showed that recasting source localization as a cosparse inverse problem allows to scale up to 3-dimensional problems which were untractable with the counterpart sparse approach. Moreover, we have confirmed that our model takes indeed advantage of reverberation, by showing that localization remains possible when the sources and the microphones are partly separated by an acoustically opaque obstacle (a situation where TDOA would obviously fail). These two results were published and presented during ICASSP'14 [37]. Recent results also include algorithmic improvements (through the use of the Alternating Direction Method of Multipliers (ADMM) framework), and evidence that, in addition to its scaling capabilities, the sparse analysis computational cost can even *benefit* from an increase in the number of measurements. A journal paper including these new results and presenting them jointly with co-space modeling in the context of brain source localization (see Section 6.6.4) is under preparation.

6.6.4. Brain source localization

Participants: Laurent Albera, Srdan Kitic, Nancy Bertin, Rémi Gribonval.

Main collaborations (tensor-based approaches): Hanna Becker (GIPSA & LTSI, France), Isabelle Merlet (LTSI, France), Fabrice Wendling (LTSI, France), Pierre Comon (GIPSA, France), Christian Benar (La Timone, Marseille), Martine Gavaret (La Timone, Marseille), Gwénaël Birot (FBML, Genève), Martin Haardt (TUI, Germany)

Main collaborations (from tensor to sparse models): Hanna Becker (GIPSA & LTSI, France), Pierre Comon (GIPSA, France), Isabelle Merlet (LTSI, France), Fabrice Wendling (LTSI, France)

Main collaborations (a sparsity-based approach): Hanna Becker (Technicolor, France), Pierre Comon (GIPSA, France), Isabelle Merlet (LTSI, France)

Main collaborations (a multimodal sparsity-based approach): Thomas Oberlin, Pierre Maurel, Christian Barillot (EPI VISAGES, Rennes, France)

Tensor-based approaches

The localization of several simultaneously active brain regions having low signal-to-noise ratios is a difficult task. To do this, tensor-based preprocessing can be applied, which consists in constructing a Space-Time-Frequency (STF) or Space-Time-Wave-Vector (STWV) tensor and decomposing it using the CP decomposition. We proposed a new algorithm for the accurate localization of extended sources based on the results of the tensor decomposition. Furthermore, we conducted a detailed study of the tensor-based preprocessing methods, including an analysis of their theoretical foundation, their computational complexity, and their performance for realistic simulated data in comparison to three conventional source localization algorithms, namely sLORETA [104], cortical LORETA (cLORETA) [103], and 4-ExSo-MUSIC [67]. Our objective consisted, on the one hand, in demonstrating the gain in performance that can be achieved by tensor-based preprocessing, and, on the other hand, in pointing out the limits and drawbacks of this method. Finally, we validated the STF and STWV techniques on real epileptic measurements to demonstrate their usefulness for practical applications. This work was published in the Elsevier NeuroImage journal [13].

From tensor to sparse models

The brain source imaging problem has been widely studied during the last decades, giving rise to an impressive number of methods using different priors. Nevertheless, a thorough study of the latter, including especially sparse and tensor-based approaches, is still missing. Consequently, we proposed i) a taxonomy of the methods based on a priori assumptions, ii) a detailed description of representative algorithms, iii) a review of identifiability results and convergence properties of different techniques, and iv) a performance comparison of the selected methods on identical data sets. Our aim was to provide a reference study in the biomedical engineering domain which may also be of interest for other areas such as wireless communications, audio source localization, and image processing where ill-posed linear inverse problems are encountered and to identify promising directions for future research in this area. A part of this work was presented at ICASSP'14 [30] while the whole part was submitted to IEEE Signal Processing Magazine.

A cosparsity-based approach

Cosparse modeling is particularly attractive when the signals of interest satisfy certain physical laws that naturally drive the choice of an analysis operator. We showed how to derive a reduced non-singular analysis operator describing EEG signals from Poisson's equation, Kirchhoff's law and some other physical constraints. As a result, we proposed the CoRE (Cosparse Representation of EEG signals) method to solve the classical brain source imaging problem. Computer simulations demonstrated the numerical performance of the CoRE method in comparison to a dictionary-based sparse approach. This work was partially presented at MLSP'14 [28].

A sparsity-based approach

Identifying the location and spatial extent of several highly correlated and simultaneously active brain sources from EEG recordings and extracting the corresponding brain signals is a challenging problem. In our comparison of source imaging techniques presented at ICASSP'14 [30], the VB-SCCD algorithm [79], which exploits the sparsity of the variational map of the sources, proved to be a promising approach. We proposed several ways to improve this method. In order to adjust the size of the estimated sources, we added a regularization term that imposes sparsity in the original source domain. Furthermore, we demonstrated the application of ADMM, which permitted to efficiently solve the optimization problem. Finally, we also considered the exploitation of the temporal structure of the data by employing L1,2-norm regularization. The performance of the resulting algorithm, called L1,2-SVB-SCCD, was evaluated based on realistic simulations in comparison to VB-SCCD and several state-of-the-art techniques for extended source localization. This work was partially presented at EUSIPCO'14 [29] and a journal paper is in preparation.

A multimodal sparsity-based approach

In the context of the HEMISFER Comin Labs project (see Section 8.1.1.1), in collaboration with the VISAGES team, we investigated brain imaging using simultaneously recorded electroencephalography (EEG) and functional magnetic resonance imaging (fMRI). To this end, we introduced a linear coupling model that links the electrical EEG signal to the hemodynamic response from the blood-oxygen level dependent (BOLD) signal. Both modalities are then symmetrically integrated, to achieve a high resolution in time and space while allowing some robustness against potential decoupling of the BOLD effect. The joint imaging problem is expressed as a linear inverse problem, which is addressed using sparse regularization. The sparsity prior naturally reflects the fact that only few areas of the brain are activated at a certain time, and it is easily implemented through proximal algorithms. At this stage, the signifiance of the method and its effectiveness have been demonstrated through numerical investigations on a simplified head model and simulated data on a realistic brain model. A conference paper has been submitted and a journal paper is in preparation.

6.6.5. Independent component analysis

Participant: Laurent Albera.

Main collaboration: Sepideh Hajipour (LTSI & BiSIPL), Isabelle Merlet (LTSI, France), Mohammad Bagher Shamsollahi (BiSIPL, Iran)

Independent Component Analysis (ICA) is a very useful tool to process biomedical signals including EEG data. We proposed a Jacobi-like Deflationary ICA algorithm, named JDICA. More particularly, while a projection-based deflation scheme inspired by Delfosse and Loubaton's ICA technique (DelL$^{\mathbb{R}}$) [78] was used, a Jacobi-like optimization strategy was proposed in order to maximize a fourth order cumulant-based contrast built from whitened observations. Experimental results obtained from simulated epileptic data mixed with a real muscular activity and from the comparison in terms of performance and numerical complexity with the FastICA [93], RobustICA [116] and DelL$^{\mathbb{R}}$ algorithms, show that the proposed algorithm offers the best trade-off between performance and numerical complexity. This work was submitted for publication in the IEEE Signal Processing Letters journal.

6.6.6. Semi-nonnegative independent component analysis

Participant: Laurent Albera.

Main collaboration: Lu Wang (LTSI, France), Amar Kachenoura (LTSI, France), Lotfi Senhadji (LTSI, France), Huazhong Shu (LIST, China)

ICA plays also an important role in many other areas including speech and audio [62], [63], [75], [72], radiocommunications [77] and document restoration [113] to cite a few.

For instance in [113], the authors use ICA to restore digital document images in order to improve the text legibility. Indeed, under the statistical independence assumption, authors succeed in separating foreground text and bleed-through/show-through in palimpsest images. Furthermore, authors in [80] use ICA to solve the ambiguity in X-ray images due to multi-object overlappings. They presented a novel object decomposition technique based on multi-energy plane radiographs. This technique selectively enhances an object that is characterized by a specific chemical composition ratio of basis materials while suppressing the other overlapping objects. Besides, in the context of classification of tissues and more particularly of brain tumors [107], ICA is very effective. In fact, it allows for feature extraction from Magnetic Resonance Spectroscopy (MRS) signals, representing them as a linear combination of tissue spectra, which are as independent as possible [111]. Moreover, using the JADE algorithm [73] applied to a mixture of sound waves computed by means of the constant-Q transform (Fourier transform with log-frequency) of a temporal waveform broken up into a set of time segments, the authors of [72] describe trills as a set of note pairs described by their spectra and corresponding time envelopes. In this case, pitch and timing of each note present in the trill can be easily deduced.

All the aforementioned applications show the high efficiency of the ICA and its robustness to the presence of noise. Despite this high efficiency in resolving the proposed applicative problems, authors did not fully exploit properties enjoyed by the mixing matrix such as its nonnegativity. For instance in [80], the thickness of each organ, which stands for the mixing coefficient, is real positive. Furthermore, reflectance indices in [113] for the background, the overwriting and the underwriting, which correspond to the mixing coefficients, are also nonnegative. Regarding tissue classification from MRS data, each observation is a linear combination of independent spectra with positive weights representing concentrations [90]; the mixing matrix is again nonnegative.

By imposing the nonnegativity of the mixing matrix within the ICA process, we shown through computer results that the extraction quality can be improved. Exploiting the nonnegativity property of the mixing matrix during the ICA process gives rise to what we call semi-nonnegative ICA. More particularly, we performed the latter by computing a constrained joint CP decomposition of cumulant arrays of different orders [99] having the nonnegative mixing matrix as loading matrices. After merging the entries of the cumulant arrays in the same third order array, the reformulated problem follows the semi-symmetric semi-nonnegative CP model defined in section 6.5.1 . Hence we use the new methods described in section 6.5.1 to perform semi-nonnegative ICA. Performance results in audio and biomedical engineering were given in the different papers cited in section 6.5.1 .

6.7. Audio and speech content processing

Audio segmentation, speech recognition, motif discovery, audio mining

6.7.1. Audio motif discovery

Participants: Frédéric Bimbot, Nathan Souviraà -Labastie.

This work was performed in close collaboration with Guillaume Gravier from the Limkmedia project-team.

As an alternative to supervised approaches for multimedia content analysis, where predefined concepts are searched for in the data, we investigate content discovery approaches where knowledge emerge from the data. Following this general philosophy, we pursued work on motif discovery in audio contents.

Audio motif discovery is the task of finding out, without any prior knowledge, all pieces of signals that repeat, eventually allowing variability. The developed algorithms allows discovering and collecting occurrences of repeating patterns in the absence of prior acoustic and linguistic knowledge, or training material.

We have designed a system to create audio thumbnails of spoken content, i.e., short audio summaries representative of the entire content, without resorting to a lexical representation. As an alternative to searching for relevant words and phrases in a transcript, unsupervised motif discovery is here used to find short, word-like, repeating fragments at the signal level without acoustic models. The output of the word discovery algorithm is exploited via a maximum motif coverage criterion to generate a thumbnail in an extractive manner. A limited number of relevant segments are chosen within the data so as to include the maximum number of motifs while remaining short enough and intelligible.

Evaluation has been performed on broadcast news reports with a panel of human listeners judging the quality of the thumbnails. Results indicate that motif-based thumbnails stand btween random thumbnails and ASR-based keywords, however still far behind thumbnails and keywords humanly authored [35].

6.7.2. Mobile device for the assistance of users in potentially dangerous situations

Participants: Romain Lebarbenchon, Ewen Camberlein, Frédéric Bimbot.

The S-Pod project is a cooperative project between industry and academia aiming at the development of mobile systems for the detection of potentially dangerous situations in the immediate environment of a user, without requiring his/her active intervention.

In this context, the PANAMA research group is involved in the design of algorithms for the analysis and monitoring of the acoustic scene around the user, yielding information which can be fused with other sources of information (physiological, contextual, etc...) in order to trigger an alarm when needed and subsequent appropriate measures.

This ongoing work is focused on the development of robust techniques for audio scene analysis, including statistical classification of audio segments into threat vs non-threat categories, and the use of spatial information to determine the location of the user with respect to the potential threat.

6.8. Music Content Processing and Music Information Retrieval

Acoustic modeling, non-negative matrix factorisation, music language modeling, music structure

6.8.1. Music structure representation and decomposition

Participants: Frédéric Bimbot, Corentin Guichaoua, Anaik Olivero.

Main collaboration: E. Vincent (EPI PAROLE, Inria Nancy), E. Deruty (external consultant)

Interest has been steadily growing in semantic audio and music information retrieval for the description of music structure, i.e. the global organization of music pieces in terms of large-scale structural units. Our group has defined a detailed methodology for the semiotic description of music structure, based on concepts and criteria which are formulated as generically as possible, i.e. without resorting to intrinsic properties of the musical content, but rather on relationships between musical elements resulting in well-identifiable patterns. The essential principles and practices developed during an annotation effort deployed by our research group on audio data, in the context of the Quaero project, has led to the public release of over 380 annotations of pop songs from three different data sets (http://musicdata.gforge.inria.fr/structureAnnotation.html) documented by a technical report which includes a few case studies and a concise statistical overview of the annotated data [31]. From the algorithmic point of view, we are currently exploring tree-based representation of music structure where a sequence of musical elements and their relationships are modeled hierarchically as the derivation of a context-free grammar. Parsimony criteria and specific cost functions adapted to music patterns are used to learn the grammar rules and the possibility of distorting the rules is introduced to account for variability accross different repetitions of musical segments.

PAREO Project-Team

6. New Results

6.1. Static Analysis

Participant: Sergueï Lenglet.

6.1.1. Static Analysis for Control Operators

Control operators, such as *call/cc* in Scheme or SML, allow programs to have access and manipulate their execution context. We study the behavioral theory of the $\lambda\mu$-calculus, an extension of the λ-calculus with a control feature similar to *call/cc*. In [6], we define an applicative bisimilarity for the $\lambda\mu$-calculus, demonstrating the differences in the definitions between call-by-name and call-by-value. We give equivalence examples to illustrate how our relations can be used; in particular, we prove David and Py's counter-example, which cannot be proved with the preexisting bisimilarities for the $\lambda\mu$-calculus. The proofs are in the accompanying research report [8], where we also define environmental bisimulations for the calculus.

6.1.2. Polymorphism and Higher-order Functions for XML

In [7], we define a calculus with higher-order polymorphic functions, recursive types with arrow and product type constructors and set-theoretic type connectives (union, intersection, and negation). We study the explicitly-typed version of the calculus in which type instantiation is driven by explicit instantiation annotations. In particular, we define an explicitly-typed λ-calculus with intersection types and an efficient evaluation model for it. The work presented in this article provides the theoretical foundations needed to design and implement higher-order polymorphic functional languages for semi-structured data.

6.2. Termination under Strategies

Participants: Horatiu Cirstea, Sergueï Lenglet, Pierre-Etienne Moreau.

Several approaches for proving the confluence and the termination of term rewriting systems have been proposed [10] and the corresponding techniques have been implemented in tools like Aprove [17] and TTT2 [26]. On the other hand, there are relatively few works on the study of these properties in the context of strategic rewriting and the corresponding results were generally obtained for some specific strategies and not within a generic framework. It would thus be interesting to reformulate these notions in the general formalism we have previously proposed [15] and to establish confluence and termination conditions similar to the ones used in standard rewriting.

We have first focused on the termination property and we targeted the rewriting strategies of the *Tom* language. We propose a direct approach which consists in translating *Tom* strategies into a rewriting system which is not guided by a given evaluation strategy and we show that our systematic transformation preserves the termination. This allowed us to take advantage of the termination proof techniques available for standard rewriting and in particular to use existing termination tools (such as Aprove and TTT2) to prove the termination of strategic rewriting systems. The efficiency and scalability of these latter tools has a direct impact on the performances of our approach especially for complex strategies for which an important number of rewrite rules could be generated. We have nevertheless proposed a meta-level implementation of the automatic transformation which improves significantly the performances of the approach. The corresponding tool is available at http://gforge.inria.fr/projects/tom.

6.3. Property-based Testing

Participants: Nauval Atmaja, Horatiu Cirstea, Pierre-Etienne Moreau.

Quality is crucial for software systems and several aspects should be taken into account. Formal verification techniques like model checking and automated theorem proving can be used to guarantee the correctness of finite or infinite systems. While these approaches provide a high level of confidence they are sometimes difficult and expensive to apply. Software testing is another approach and although it cannot guarantee correctness it can be very efficient in finding errors.

We have proposed a property based testing framework for the *Tom* language inspired from the ones proposed in the context of functional programming. The previously developed tool has been improved by integrating it in the *Junit* framework. The tests are thus highly automatized and the library can be smoothly integrated in classical IDEs. The relatively simple shrinking method which searches a smaller counter-example starting from an initial relatively complex one has been also improved. The library is available at http://gforge.inria.fr/projects/tom.

6.4. Inductive Reasoning

Participant: Sorin Stratulat.

6.4.1. Decision Procedures to Prove Inductive Theorems Without Induction

Automated inductive reasoning for term rewriting has been extensively studied in the literature. Classes of equations and term rewriting systems (TRSs) with decidable inductive validity have been identified and used to automatize the inductive reasoning. In [9], we give procedures for deciding the inductive validity of equations in some standard TRSs on natural numbers and lists. Contrary to previous decidability results, our procedures can automatically decide without involving induction reasoning the inductive validity of arbitrary equations for these TRSs, that is, without imposing any syntactical restrictions on the form of equations. We also report on the complexity of our decision procedures. These decision procedures are implemented in our automated provers for inductive theorems of TRSs and experiments are reported.

6.4.2. Implementing Reasoning Modules in Implicit Induction Theorem Provers

In [30], we detail the integration in SPIKE, an implicit induction theorem prover, of two reasoning modules operating over naturals combined with interpreted symbols. The first integration schema is à la Boyer-Moore, based on the combination of a congruence closure procedure with a decision procedure for linear arithmetic over rationals/reals. The second follows a 'black-box' approach and is based on external SMT solvers. It is shown that the two extensions significantly increase the power of SPIKE; their performances are compared when proving a non-trivial application.

6.4.3. Building Explicit Induction Schemas for Cyclic Induction Reasoning

In the setting of classical first-order logic with inductive predicates, two kinds of sequent-based induction reasoning are distinguished: cyclic and structural. Proving their equivalence is of great theoretical and practical interest for the automated reasoning community. Previously, it has been shown how to transform any structural proof developed with the LKID system into a cyclic proof using the CLKID$^\omega$ system. However, the inverse transformation was only conjectured. In [29], we provide a simple procedure that performs the inverse transformation for an extension of LKID with explicit induction rules issued from the structural analysis of CLKID$^\omega$ proofs, then establish the equivalence of the two systems. This result is further refined for an extension of LKID with Noetherian induction rules. We show that Noetherian induction subsumes the two kinds of reasoning. This opens the perspective for building new effective induction proof methods and validation techniques supported by (higher-order) certification systems integrating the Noetherian induction principle, like Coq.

PARIETAL Project-Team

6. New Results

6.1. Highlights of the Year

- Congratulations also to Alex and Daniel Strohmeier for their best paper award at the PRNI 2014 conference: “Improved MEG/EEG source localization with reweighted mixed-norms“.
- Elvis Dohmatob got a honorable mention for the student paper award at PRNI 2014 for the work “Benchmarking solvers for TV-l1 least-squares and logistic regression in brain imaging”

6.2. Which fMRI clustering gives good brain parcellations?

Participants: Bertrand Thirion [Correspondant], Gaël Varoquaux, Elvis Dohmatob.

Analysis and interpretation of neuroimaging data often require one to divide the brain into a number of regions, or parcels, with homogeneous characteristics, be these regions defined in the brain volume or on on the cortical surface. While predefined brain atlases do not adapt to the signal in the individual subjects images, parcellation approaches use brain activity (e.g. found in some functional contrasts of interest) and clustering techniques to define regions with some degree of signal homogeneity. In this work, we address the question of which clustering technique is appropriate and how to optimize the corresponding model. We use two principled criteria: goodness of fit (accuracy), and reproducibility of the parcellation across bootstrap samples. We study these criteria on both simulated and two task-based functional Magnetic Resonance Imaging datasets for the Ward, spectral and K-means clustering algorithms. We show that in general Ward's clustering performs better than alternative methods with regard to reproducibility and accuracy and that the two criteria diverge regarding the preferred models (reproducibility leading to more conservative solutions), thus deferring the practical decision to a higher level alternative, namely the choice of a trade-off between accuracy and stability.

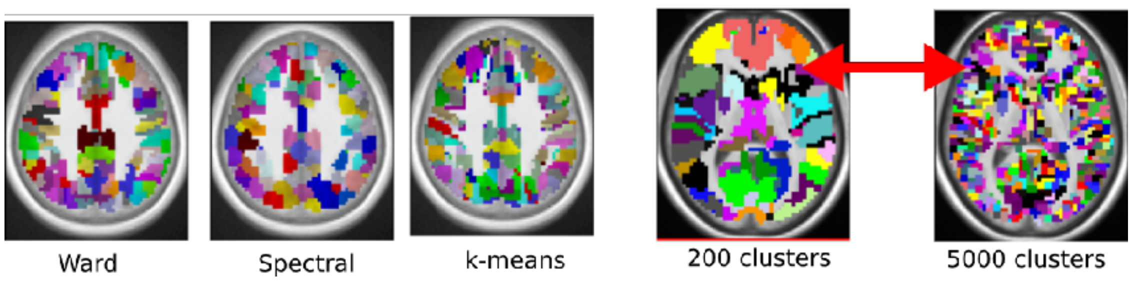


Figure 3. Practitioner have to decide which clustering method to use and how to select the number of clusters. In [21], we provide empirical guidelines and criteria to guide that choice in the context of functional brain imaging.

More details can be found in [21].

6.3. Principal Component Regression predicts functional responses across individuals

Participants: Bertrand Thirion [Correspondant], Gaël Varoquaux, Olivier Grisel.

Inter-subject variability is a major hurdle for neuroimaging group-level inference, as it creates complex image patterns that are not captured by standard analysis models and jeopardizes the sensitivity of statistical procedures. A solution to this problem is to model random subjects effects by using the redundant information conveyed by multiple imaging contrasts. In this paper, we introduce a novel analysis framework, where we estimate the amount of variance that is fit by a random effects subspace learned on other images; we show that a principal component regression estimator outperforms other regression models and that it fits a significant proportion (10% to 25%) of the between-subject variability. This proves for the first time that the accumulation of contrasts in each individual can provide the basis for more sensitive neuroimaging group analyzes.

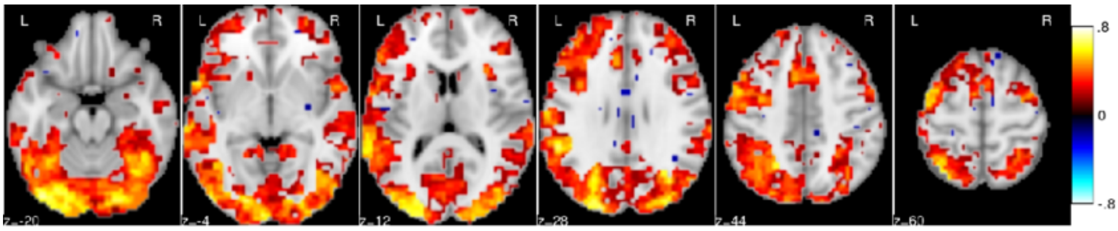


Figure 4. In most brain regions, knowing the amount of activation related to a set of reference contrasts yields an accurate prediction of the activation for a target contrast.

More details can be found in [36].

6.4. Deriving a multi-subject functional-connectivity atlas to inform connectome estimation

Participants: Ronald Phlypo [Correspondant], Bertrand Thirion, Gaël Varoquaux.

The estimation of functional connectivity structure from functional neuroimaging data is an important step toward understanding the mechanisms of various brain diseases and building relevant biomarkers. Yet, such inferences have to deal with the low signal-to-noise ratio and the paucity of the data. With at our disposal a steadily growing volume of publicly available neuroimaging data, it is however possible to improve the estimation procedures involved in connectome mapping. In this work, we propose a novel learning scheme for functional connectivity based on sparse Gaussian graphical models that aims at minimizing the bias induced by the regularization used in the estimation, by carefully separating the estimation of the model support from the coefficients. Moreover, our strategy makes it possible to include new data with a limited computational cost. We illustrate the physiological relevance of the learned prior, that can be identified as a functional connectivity atlas, based on an experiment on 46 subjects of the Human Connectome Dataset.

More details can be found in [35].

6.5. Machine Learning Patterns for Neuroimaging-Genetic Studies in the Cloud

Participants: Virgile Fritsch, Bertrand Thirion, Gaël Varoquaux.

Brain imaging is a natural intermediate phenotype to understand the link between genetic information and behavior or brain pathologies risk factors. Massive efforts have been made in the last few years to acquire high-dimensional neuroimaging and genetic data on large cohorts of subjects. The statistical analysis of such data is carried out with increasingly sophisticated techniques and represents a great computational challenge. Fortunately, increasing computational power in distributed architectures can be harnessed, if new neuroinformatics infrastructures are designed and training to use these new tools is provided. Combining a

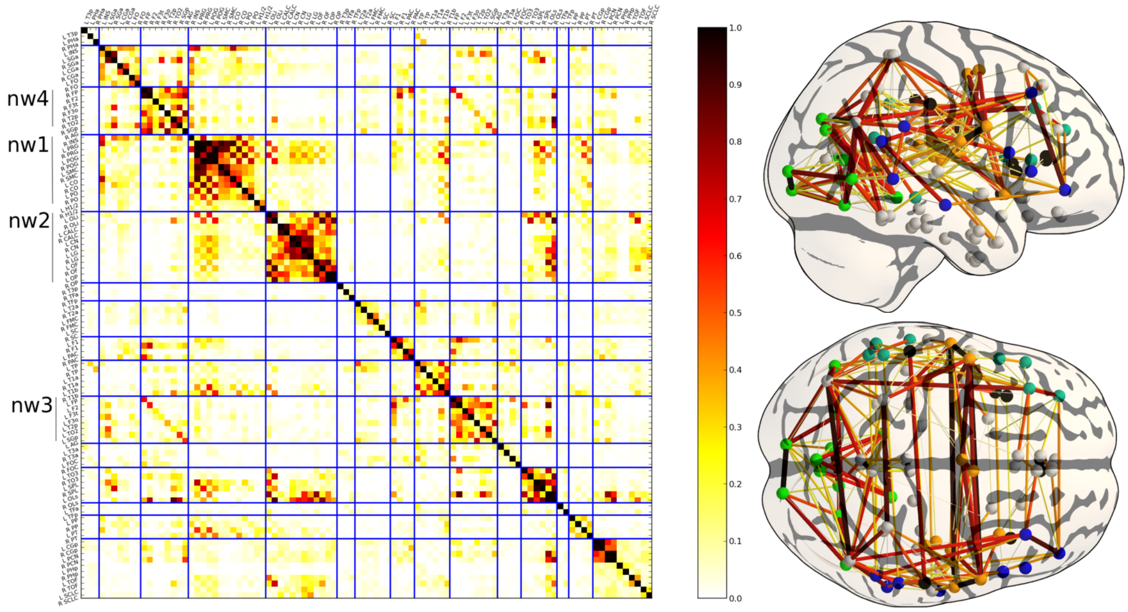


Figure 5. Prior on the functional connectivity: the coefficient of the matrix represent the frequency of an edge at each position. This model can be interpreted as a data-driven atlas of brain functional connections. In the current framework, it can easily be updated to take into account more data.

MapReduce framework (TomusBLOB) with machine learning algorithms (Scikit-learn library), we design a scalable analysis tool that can deal with non-parametric statistics on high-dimensional data. End-users describe the statistical procedure to perform and can then test the model on their own computers before running the very same code in the cloud at a larger scale. We illustrate the potential of our approach on real data with an experiment showing how the functional signal in subcortical brain regions can be significantly fit with genome-wide genotypes. This experiment demonstrates the scalability and the reliability of our framework in the cloud with a two weeks deployment on hundreds of virtual machines.

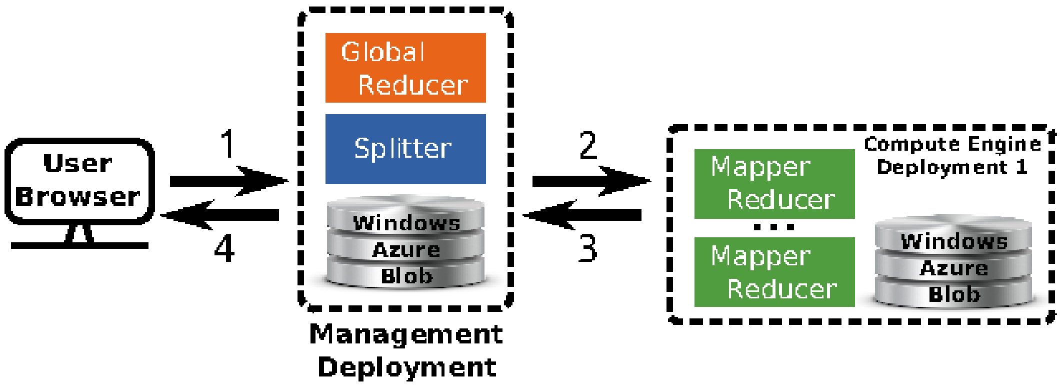


Figure 6. Overview of the multi site deployment of a hierarchical Tomus-MapReduce compute engine. 1) The end-user uploads the data and configures the statistical inference procedure on a webpage. 2) The Splitter partitions the data and manages the workload. The compute engines retrieves job information trough the Windows Azure Queues. 3) Compute engines perform the map and reduce jobs. The management deployment is informed of the progression via the Windows Azure Queues system and thus can manage the execution of the global reducer. 4) The user downloads the results of the computation on the webpage of the experiment.

More details can be found in [17].

6.6. Data-driven HRF estimation for encoding and decoding models

Participants: Fabian Pedregosa Izquierdo [correspondant], Michael Eickenberg, Alexandre Gramfort, Philippe Ciuciu, Bertrand Thirion, Gaël Varoquaux.

Despite the common usage of a canonical, data-independent, hemodynamic response function (HRF), it is known that the shape of the HRF varies across brain regions and subjects. This suggests that a data-driven estimation of this function could lead to more statistical power when modeling BOLD fMRI data. However, unconstrained estimation of the HRF can yield highly unstable results when the number of free parameters is large. We develop a method for the joint estimation of activation and HRF using a rank constraint causing the estimated HRF to be equal across events/conditions, yet permitting it to be different across voxels. Model estimation leads to an optimization problem that we propose to solve with an efficient quasi-Newton method exploiting fast gradient computations. This model, called GLM with Rank-1 constraint (R1-GLM), can be extended to the setting of GLM with separate designs which has been shown to improve decoding accuracy in brain activity decoding experiments. We compare 10 different HRF modeling methods in terms of encoding and decoding score in two different datasets. Our results show that the R1-GLM model significantly outperforms competing methods in both encoding and decoding settings, positioning it as an attractive method both from the points of view of accuracy and computational efficiency.

More details can be found in [19].

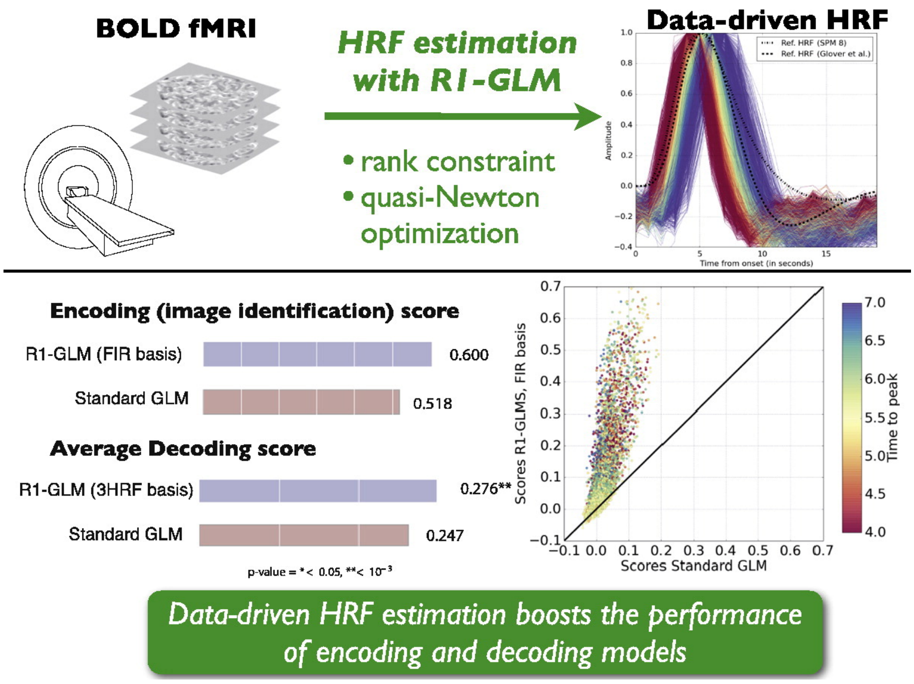


Figure 7. Illustration of the hemodynamic response function estimation framework introduced in [19].

6.7. Benchmarking solvers for TV-l1 least-squares and logistic regression in brain imaging

Participants: Elvis Dohmatob [correspondant], Michael Eickenberg, Gaël Varoquaux, Bertrand Thirion.

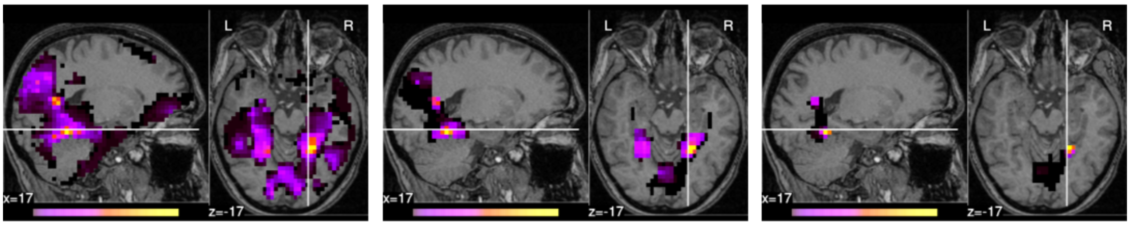
Learning predictive models from brain imaging data, as in decoding cognitive states from fMRI (functional Magnetic Resonance Imaging), is typically an ill-posed problem as it entails estimating many more parameters than available sample points. This estimation problem thus requires regularization. Total variation regularization, combined with sparse models, has been shown to yield good predictive performance, as well as stable and interpretable maps. However, the corresponding optimization problem is very challenging: it is non-smooth, non-separable and heavily ill-conditioned. For the penalty to fully exercise its structuring effect on the maps, this optimization problem must be solved to a good tolerance, resulting in a computational challenge. In this work, we explore a wide variety of solvers and exhibit their convergence properties on fMRI data. We introduce a variant of smooth solvers and show that it is a promising approach in these settings. Our findings show that care must be taken in solving TV-l1 estimation in brain imaging and highlight the successful strategies.

More details can be found in [30]

6.8. Interplay between functional connectivity and scale-free dynamics in intrinsic fMRI networks

Participant: Philippe Ciuciu [correspondant].

Studies employing functional connectivity-type analyses have established that spontaneous fluctuations in functional magnetic resonance imaging (fMRI) signals are organized within large-scale brain networks. Meanwhile, fMRI signals have been shown to exhibit 1/f-type power spectra – a hallmark of scale-free



ΔE = .1 ΔE = .001 ΔE = 10^{-5}

Figure 8. TV-l1 maps for a face-house discrimination task taken from a visual recognition dataset, with regularization parameters chosen by cross-validation, for different stopping criteria. Note that the stopping criterion is defined as a threshold on the energy decrease per one iteration of the algorithm. This figure shows the importance of convergence of the multivariate estimator, and motivates the need for a fast solver.

dynamics. We studied the interplay between functional connectivity and scale-free dynamics in fMRI signals, utilizing the fractal connectivity framework – a multivariate extension of the univariate fractional Gaussian noise model, which relies on a wavelet formulation for robust parameter estimation. We applied this framework to fMRI data acquired from healthy young adults at rest and performing a visual detection task. First, we found that scale-invariance existed beyond univariate dynamics, being present also in bivariate cross-temporal dynamics. Second, we observed that frequencies within the scale-free range do not contribute evenly to inter-regional connectivity, with a systematically stronger contribution of the lowest frequencies, both at rest and during task. Third, in addition to a decrease of the Hurst exponent and inter-regional correlations, task performance modified cross-temporal dynamics, inducing a larger contribution of the highest frequencies within the scale-free range to global correlation.

More details can be found in [16].

6.9. Supramodal processing optimizes visual perceptual learning and plasticity

Participants: Philippe Ciuciu [correspondant], Alexandre Gramfort.

Multisensory interactions are ubiquitous in cortex and it has been suggested that sensory cortices may be supramodal i.e. capable of functional selectivity irrespective of the sensory modality of inputs. Here, we asked whether learning to discriminate visual coherence could benefit from supramodal processing. To this end, three groups of participants were briefly trained to discriminate which of a red or green intermixed population of random-dot-kinematograms (RDKs) was most coherent in a visual display while being recorded with magnetoencephalography (MEG). During training, participants heard no sound (V), congruent acoustic textures (AV) or auditory noise (AVn); importantly, congruent acoustic textures shared the temporal statistics – i.e. coherence – of visual RDKs. After training, the AV group significantly outperformed participants trained in V and AVn although they were not aware of their progress. In pre- and post-training blocks, all participants were tested without sound and with the same set of RDKs. When contrasting MEG data collected in these experimental blocks, selective differences were observed in the dynamic pattern and the cortical loci responsive to visual RDKs. First and common to all three groups, vlPFC showed selectivity to the learned coherence levels whereas selectivity in visual motion area hMT+ was only seen for the AV group. Second and solely for the AV group, activity in multisensory cortices (mSTS, pSTS) correlated with post-training performances; additionally, the latencies of these effects suggested feedback from vlPFC to hMT+ possibly mediated by temporal cortices in AV and AVn groups. Altogether, we interpret our results in the context of the Reverse Hierarchy Theory of learning in which supramodal processing optimizes visual perceptual learning by capitalizing on sensory-invariant representations - here, global coherence levels across sensory modalities.

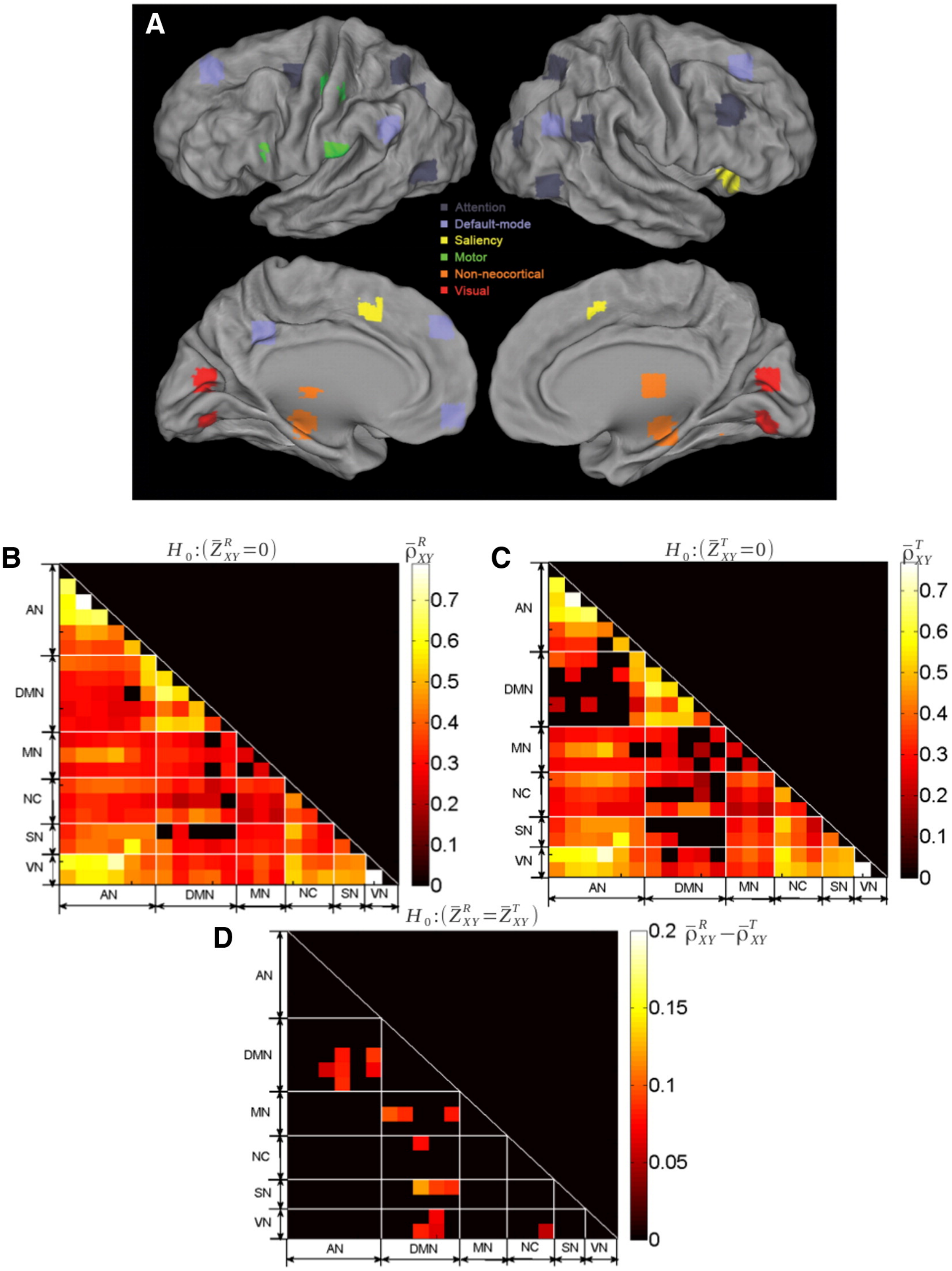


Figure 9. Networks definition and correlation structure. Top (A): ROIs mapped onto the cortical surface, with each color denoting a different network. Middle (B): Group-averaged inter-regional correlation matrix at rest ($p < 0.05$, Bonferroni corrected). Regions are grouped by network to ease visualization. Middle (C): Group-averaged inter-regional correlation matrix during the visual detection task ($p < 0.05$, Bonferroni corrected). Bottom (D): Difference of the correlation coefficients between rest and task (thresholded at $p < 0.01$, uncorrected, two-sample t-test for rest vs. task). The ROIs are grouped by networks; these networks correspond to the diagonal triangles surrounded by white dashed lines.

More details can be found in [25].

6.10. Variable density sampling with continuous trajectories. Application to MRI.

Participants: Nicolas Chauffert, Philippe Ciuciu [correspondant].

Reducing acquisition time is a crucial challenge for many imaging techniques. Compressed Sensing (CS) theory offers an appealing framework to address this issue since it provides theoretical guarantees on the reconstruction of sparse signals by projection on a low dimensional linear subspace. In this paper, we focus on a setting where the imaging device allows to sense a fixed set of measurements. We first discuss the choice of an optimal sampling subspace (smallest subset) allowing perfect reconstruction of sparse signals. Its standard design relies on the random drawing of independent measurements. We discuss how to select the drawing distribution and show that a mixed strategy involving partial deterministic sampling and independent drawings can help breaking the so-called "coherence barrier". Unfortunately, independent random sampling is irrelevant for many acquisition devices owing to acquisition constraints. To overcome this limitation, the notion of Variable Density Samplers (VDS) is introduced and defined as a stochastic process with a prescribed limit empirical measure. It encompasses samplers based on independent measurements or continuous curves. The latter are crucial to extend CS results to actual applications. Our main contribution lies in two original continuous VDS. The first one relies on random walks over the acquisition space whereas the second one is heuristically driven and rests on the approximate solution of a Traveling Salesman Problem. Theoretical analysis and retrospective CS simulations in magnetic resonance imaging highlight that the TSP-based solution provides improved reconstructed images in terms of signal-to-noise ratio compared to standard sampling schemes (spiral, radial, 3D iid...).

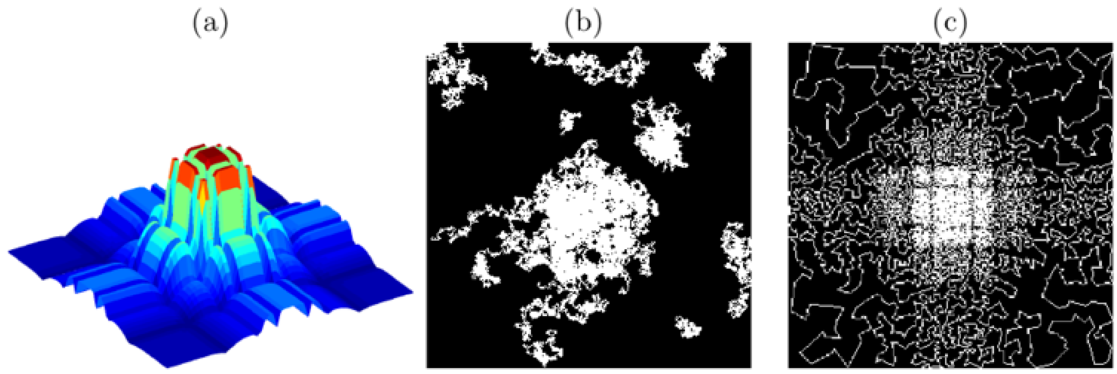


Figure 10. (a): Target distribution π to be approximated. Continuous random trajectories reaching distribution π based on Markov chains (b) and on a TSP solution (c). The latter is much more accurate.

More details can be found in [15].

PARKAS Project-Team

6. New Results

6.1. Highlights of the Year

The paper *ReactiveML, a reactive extension to ML* of Mandel and Pouzet has been declared to be the *most influential paper of PPDP (Principles and Practice of Declaractive Programming) 2005*. A previous version of the paper, submitted to JFLA'05, has been declared to be "une contribution marquante parmi les articles publiés aux JFLA".

6.2. Quasi-synchrony

Participants: Guillaume Baudart, Timothy Bourke, Marc Pouzet.

We study the implementation of critical control applications on the so-called *quasi-periodic* distributed architectures. These architectures, used in civil avionics (e.g., Airbus A380), consist of a collection of distributed processors running with *quasi-periodic* clocks, that is, un-synchronized physical clocks subject to bounded jitterring. The theory of quasi-synchrony has been introduced by Paul Caspi in the 2000' [29]. Loosely Time-Triggered Architectures (LTTA) denotes such architectures with the prototocol used to implement a synchronous program on top of it.

Over the last ten year two protocols were considered: (1) *Back-Pressure* LTTA [25] based on a acknowledgement mechanism reminiscent of elastic circuit [43]. (2) *Time-Based* LTTA [28] which uses timing constraints of the architecture to mimic a synchronous execution.

During year 2014, we have entirelly reformulated the model of LTTA using synchronous semantics and principles. Compared to previous formalizations based on Petri nets [24], this new presentation is is simpler and more uniform with the same theoretical model used for both the application and the protocol ((1) or (2)). Moreover, it is easier to consider mixed protocols (a whole application with part based on time-based communication and others based on back-pressure). Besides this, we also proposed a new and more flexible Time-Based LTTA, allowing for pipelining by not reconstructing global synchronization, unlike what was done in previous Time-Based LTTA.

6.3. Hybrid Synchronous Languages

Participants: Guillaume Baudart, Timothy Bourke, Marc Pouzet.

During year 2014, we mainly worked on two directions: (a) the design and implementation of causality analysis for hybrid systems modelers; (b) the design and implementation of a new compilation technique producing imperative sequential code.

This research is conducted in collaboration with Albert Benveniste and Benoit Caillaud (Hycomes team at Inria, Rennes), Jean-Louis Colaco, Cédric Pasteur and Bruno Pagano from the SCADE core team of Esterel-Technologies/ANSYS.

Causality analysis In this work, we address the static detection of causality loops for a hybrid modeling language that combines synchronous Lustre-like data-flow equations with Ordinary Differential Equations (ODEs). We introduce the operator *last(x)* for the left-limit of a signal x. This operator is used to break causality loops and permits a uniform treatment of discrete and continuous state variables. The semantics relies on non-standard analysis, defining an execution as a sequence of infinitesimally small steps. A signal is deemed *causally correct* when it can be computed sequentially and only progresses by infinitesimal steps outside of discrete events. The causality analysis takes the form of a simple type system. In well-typed programs, signals are proved continuous during integration.

This analysis has been presented at [4] and is fully implemented in the hybrid synchronous language Zélus.

A Synchronous-based Code Generator For Explicit Hybrid Systems Languages The generation of sequential code is important for simulations to be efficient and to produce target embedded code. While sequential code generation in hybrid modeling tools is routinely used for efficient simulation, it is little or not used for producing target embedded code in critical applications submitted to strong safety requirements. This is a break in the development chain: parts of the applications must be rewritten into either sequential or synchronous programs, and all properties verified on the source model cannot be trusted and have to be re-verified on the target code.

In this work, we present a novel approach for the code generation of a hybrid systems modeling language. By building on top of an existing synchronous language and compiler, it reuses almost all the existing infrastructure with only a few modifications. Starting from an existing synchronous data-flow language extended with Ordinary Differential Equations (ODEs), we detail the translation to sequential code. The translation is expressed as a sequence of source-to-source transformations. A generic intermediate language is introduced to represent transition functions which are turned into C code. The versatility of the compiler organisation is illustrated by considering two classical targets: generation of simulation code complying with the FMI standard and linking with an off-the-shelf numerical solver (Sundials CVODE).

This new code generation has been implemented in two different compilers: the Zélus research prototype and the industrial SCADE Suite KCG code generator, at Esterel-Technologies/ANSYS. Here, SCADE is conservatively extended with ODEs, following previous works by Benveniste et al. and implemented in Zélus. In the SCADE compiler, it was possible to reuse almost all the existing infrastructure like static checking, intermediate languages, and optimisations, with few modifications. The extension to account for hybrid features represents only 5% additional lines of code, which is surprisingly low. Moreover, the proposed language extension is conservative in that regular synchronous functions are compiled as before—the same synchronous code is used both for simulation and for execution on target platforms.

This full-scale validation confirm the interest in building a hybrid systems modeler on top of a existing synchronous language. Moreover, the precise definition of code generation, built on a proven compiler infrastructure of a synchronous language avoids the rewriting of control software and may also increase the confidence in what is simulated.

This work will be presented at the *International Conference on Compiler Construction (CC)*, in April 2015.

6.4. Fidelity in Real-Time Programming

Participants: Guillaume Baudart, Timothy Bourke.

Synchronous languages are a rigorous approach to programming, analyzing, and implementing embedded systems. Real-time aspects are typically handled by discretizing time using either (implicit) ticks or (explicit) named signals, and later verifying that the (necessarily bounded) execution time of a reaction is strictly less than the period of the fastest timing signal. This approach has many advantages: it separates logical behaviour from implementation concerns, yields a simple and precise programming model, and abstracts from eventual run-time environments. For an important subclass of embedded protocols and controllers, however, we believe it advantageous to add constructions that deal more concretely with real-time constraints.

We are pursuing these ideas in the enriched timing model provided by the Zélus programming language (detailed elsewhere). We continue to study the extension and application of this language to the modelling, simulation, analysis, and implementation of real-time embedded software.

This year we developed three case studies: quasi-synchronous architectures (from last year), loosely time-triggered architectures (detailed elsewhere), and a small embedded controller. These case studies motivate and drive our research and implicitly define the subclass of embedded systems that we aim to treat. They have each been modelled in Zélus and can be simulated with the existing compiler.

We made progress on defining a subset of Zélus that is ammenable to discretization techniques for more flexible simulation. A first version of an appropriate algorithm has been sketched and partially implemented. Work continues on developing it with the idea of incorporating it into the Zélus compiler and using it to treat our case studies.

6.5. Mechanization of AODV loop freedom proof

Participant: Timothy Bourke.

The Ad hoc On demand Distance Vector (AODV) routing protocol is described in RFC3561. It allows the nodes in a Mobile Ad hoc Network (MANET) to know where to forward messages so that they eventually reach their destinations. The nodes of such networks are *reactive systems* that cooperate to provide a global service (the sending of messages from node to node) satisfying certain correctness properties (namely 'loop freedom'—that messages are never sent in circles).

This year I finalized both the framework for network invariant proofs [20] and its application to the AODV protocol [21] and submitted them for inclusion in the *Archive of Formal Proof*, an online and open-source repository of formal developements in the Isabelle proof assistant (indexed as a journal). I presented results on the framework at the Vienna 'Summer of Logic' [6] and my colleagues presented the application in Sydney [5]. Together with an intern at NICTA and Sydney, my colleagues and I made preliminary investigations into extending the framework and model with timing details. A journal version of the ITP paper has been submitted.

In collaboration with Peter Höfner (NICTA) and Robert J. van Glabbeek (UNSW/NICTA).

6.6. Reasoning about C11 Program Transformations

Participants: Francesco Zappa Nardelli, Thibaut Balabonski, Robin Morisset.

We have shown that the weak memory model introduced by the 2011 C and C++ standards does not permit many of common source-to-source program transformations (such as expression linearisation and "roach motel" reordering) that modern compilers perform and that are deemed to be correct. As such it cannot be used to define the semantics of intermediate languages of compilers, as, for instance, LLVM aimed to. We consider a number of possible local fixes, some strengthening and some weakening the model. We have evaluated the proposed fixes by determining which program transformations are valid with respect to each of the patched models. We have provided formal Coq proofs of their correctness or counterexamples as appropriate.

A paper on this work has been accepted in [13]. In collaboration with Viktor Vafeiadis (MPI-SWS, Germany).

6.7. Language design on top of JavaScript

Participant: Francesco Zappa Nardelli.

This research project aims at improving the design of the JavaScript language. In [22] we propose a typed extension of JavaScript combining dynamic types, concrete types and like types to let developers pick the level of guarantee that is appropriate for their code. We have implemented our type system and we have explored the performance and software engineering benefits.

With Gregor Richards and Jan Vitek (Purdue University).

6.8. Tiling for Stencils

Participants: Tobias Grosser, Sven Verdoolaege, Albert Cohen.

This research project aims with optimizing time-iterated stencil operations.

Iterative stencil computations are important in scientific computing and more and more also in the embedded and mobile domain. Recent publications have shown that tiling schemes that ensure concurrent start provide efficient ways to execute these kernels. Diamond tiling and hybrid-hexagonal tiling are two tiling schemes that enable concurrent start. Both have different advantages: diamond tiling has been integrated in a general purpose optimization framework and uses a cost function to choose among tiling hyperplanes, whereas the greater flexibility with tile sizes for hybrid-hexagonal tiling has been exploited for effective generation of GPU code.

We undertook a comparative study of these two tiling approaches and proposed a hybrid approach that combines them. We analyzed the effects of tile size and wavefront choices on tile-level parallelism, and formulate constraints for optimal diamond tile shapes. We then extended, for the case of two dimensions, the diamond tiling formulation into a hexagonal tiling one, which offers both the flexibility of hexagonal tiling and the generality of the original diamond tiling implementation. We also showed how to compute tile sizes that maximize the compute-to-communication ratio, and apply this result to compare the best achievable ratio and the associated synchronization overhead for diamond and hexagonal tiling.

One particularly exciting result is the ability to apply tiling to periodic data domains. These computations are prevalent in computational sciences, particularly in partial differential equation solvers. We proposed a fully automatic technique suitable for implementation in a compiler or in a domain-specific code generator for such computations. Dependence patterns on periodic data domains prevent existing algorithms from finding tiling opportunities. Our approach augments a state-of-the-art parallelization and locality-enhancing algorithm from the polyhedral framework to allow time-tiling of stencil computations on periodic domains. Experimental results on the swim SPEC CPU2000fp benchmark show a speedup of $5\times$ and $4.2\times$ over the highest SPEC performance achieved by native compilers on Intel Xeon and AMD Opteron multicore SMP systems, respectively. On other representative stencil computations, our scheme provides performance similar to that achieved with no periodicity, and a very high speedup is obtained over the native compiler. We also report a mean speedup of about $1.5\times$ over a domain-specific stencil compiler supporting limited cases of periodic boundary conditions. To the best of our knowledge, it has been infeasible to manually reproduce such optimizations on swim or any other periodic stencil, especially on a data grid of two-dimensions or higher.

These works resulted in a number of high-profile publications, including a nommination for a best paper award, and culminated with the PhD thesis defense of Tobias Grosser.

6.9. Portable representation for polyhedral compilation

Participants: Riyadh Baghdadi, Michael Kruse, Chandan Reddy, Tobias Grosser, Sven Verdoolaege, Albert Cohen.

Programming accelerators such as GPUs with low-level APIs and languages such as OpenCL and CUDA is difficult, error prone, and not performance-portable. Automatic parallelization and domain specific languages have been proposed to hide this complexity and to regain some performance portability. We proposed PENCIL, a subset of GNU C99 with specific programming rules. A compiler for a Domain-Specific Language (DSL) may use it as a target language, a domain expert may use it as a portable implementation language facilitating the parallelization of real-world applications, and an optimization expert may use PENCIL to accelerate legacy applications.

The design of PENCIL is simultaneously a key research result and a milestone for parallelizing compiler engineering/design. Aspects of its static-analysis-friendly, formal semantics are highly original, for the language's ability to preserve expressiveness and modularity without jeopardizing a (polyhedral) compiler's ability to perform aggressive transformations. We validated its potential as a front-end to a state-of-the-art polyhedral compiler, extending its applicability to dynamic, data dependent control flow and non-affine array accesses. We illustrated this PENCIL-enabled flow on the generation of highly optimized OpenCL code, considering a set of standard benchmarks (Rodinia and SHOC), image processing kernels, and DSL embedding scenarios for linear algebra (BLAS) and signal processing radar applications (SPEAR-DE). We ran experimental results on a variety of platforms, including an AMD Radeon HD 5670 GPU, an Nvidia GTX470 GPU, and an ARM Mali-T604 GPU.

This work is conducted in collaboration with partners from ARM, RealEyes (a computer vision company) and Imperial College.

6.10. Correct and efficient runtime systems

Participants: Nhat Minh Lê, Robin Morisset, Adrien Guatto, Albert Cohen.

Complementing our different compilation efforts for synchronous and task-parallel data-flow languages, we studied the implementation of Kahn process networks, a deterministic parallel programming model, on shared memory multiprocessors. This model is based on a familiar abstraction: blocking communication through bounded, in-order, single-producer single-consumer queues.

We proposed two novel algorithms that construct such blocking queues on top of concurrent ring buffers and user-land scheduling components. We implemented our algorithms in C11, taking advantage of the relaxed memory model of the language, and prove the correctness of this implementation.

We used these algorithms in a complete runtime system for Kahn process networks with applications ranging from linear algebra kernels to stream computing. In particular, our implementations of the Cholesky and LU factorizations outperform state-of-the-art parallel linear algebra libraries on commodity x86 hardware.

6.11. A Functional Synchronous Language with Integer Clocks

Participants: Adrien Guatto, Albert Cohen, Louis Mandel, Marc Pouzet.

Synchronous languages in the vein of Lustre are first-order functional languages dedicated to stream processing. Lustre compilers use a type-like static analysis, the clock calculus, to reject programs that cannot be implemented as finite state machines. The broad idea is to assign to each element of a stream a logical computation date in a global, discrete time scale. When this analysis succeeds, the types obtained guide the code generation phase of the compiler, which produces transition functions. In practice, these functions consists in simple, bounded memory C code featuring only assignments and conditional statements.

This research work explores a variation on Lustre and its compilation. Our proposal is twofold. First, we add a new construct that creates a local time scale whose internal steps are invisible from the outside. Second, we change the clock calculus to allow several elements of a stream to be computed during the same time step. The resulting type system comes with a soundness proof, which relies on an elementary form of step-indexed realizability, and with a code generation scheme adapted to the new setting, and featuring nested loops in the target code.

PARSIFAL Project-Team

6. New Results

6.1. Highlights of the Year

Dale Miller's 1994 LICS paper titled "A Multiple-Conclusion Meta-Logic" [67] was a co-recipient of the LICS Test of Time Award.

6.2. Modular Systems for Classical and Intuitionistic Logic

Participants: Sonia Marin, Lutz Straßburger.

Last year we have shown deductive systems for all intuitionistic modal logics in the modal S5-cube using logical rules in nested sequents [75]. This year we managed to exhibit fully modular systems. That is to say that there is a bijective correspondence between the modal axioms and the inference rules in the deductive system. This is achieved by using a combination of structural and logical rules. This result has been presented at AiML 2014 [24].

6.3. Nested Sequents for Constructive Modal Logics

Participants: Ryuta Arisaka, Anupam Das, Lutz Straßburger.

In the propositional case, "constructive" and "intuitionistic" logic are usually considered the same. However, in the presence of the modalities $\Box$ and $\diamond$ this situation changes because there are several choice of which variants of the k-axiom (which are all equivalent in the classical case) are to be included. Whereas in [75] the intuitionistic variant of the S5-cube has been studied, we studied in this years work [34] the constructive variant of the logics in the S5-cube.

6.4. Intuitionistic Logic in the Calculus of Structures

Participants: Nicolas Guenot, Lutz Straßburger.

The calculus of structures has mainly be used for "classical" logics that come with a De Morgan duality. The reason is that all normalization procedures developed so far for the calculus of structures rely on this De Morgan duality.

In this work, we give two proof systems for implication-only intuitionistic logic in the calculus of structures. The first is a direct adaptation of the standard sequent calculus to the deep inference setting. It comes with a cut elimination procedure that is similar to the one from the sequent calculus, using a non-local rewriting. The second system is the symmetric completion of the first, as normally given in deep inference for logics with a De Morgan duality: all inference rules have duals, as cut is dual to the identity axiom. For this symmetric system we prove a generalization of cut elimination, that we call symmetric normalization, where all rules dual to standard ones are permuted up in the derivation. The result is a decomposition theorem having cut elimination and interpolation as corollaries. This work has been presented at the CSL-LICS 2014 conference [22].

6.5. Free Theorems for Curry

Participant: Lutz Straßburger.

Free theorems [79] are a means of type-based reasoning and are being successfully applied for typed functional programming languages like Haskell, e.g., for program transformation and generally establishing semantic properties [53], [78]. As a simple example, for every polymorphic function $f :: [\alpha] \to [\alpha]$ from lists to lists, arbitrary types τ_1 and τ_2, and a function $g :: \tau_1 \to \tau_2$, we have $f \circ (\mathtt{map}\ g) = (\mathtt{map}\ g) \circ f$, for the standard function $\mathtt{map} :: (\alpha \to \beta) \to [\alpha] \to [\beta]$ which takes a function and a list and applies that function to every entry of the list. It would be of interest to also have such free theorems available for typed functional-logic languages like Curry.

Previous work [48] has investigated free theorems for such a language, Curry [60], phenomenologically and provides intuition for premises of free theorems as well as counterexamples. Proof of the positive claims has been elusive so far, mainly because Curry's type system fails to reflect the key feature: nondeterminism. This avoidance is convenient for programmers, as they do not have to distinguish between deterministic and nondeterministic values. However, it is a hindrance to formal reasoning: the conditions identified in [48] include a notion of determinism, and hence it is a serious weakness of the type system not to capture this.

In a joint work with colleagues at the University of Bonn, published in [25], we have developed an intermediate language, called SaLT, that allowed us to prove a *Parametricity Theorem* which could be used to derive free theorems for Curry.

This work is the result of the PHC Procope collaboration with the University of Bonn (duration 2012-2013).

6.6. A logical basis for quantum evolution and entanglement

Participant: Lutz Straßburger.

In discrete quantum causal dynamics, quantum systems are viewed as discrete structures, namely directed acyclic graphs. In such a graph, events are considered as vertices and edges depict propagation between events. Evolution is described as happening between a special family of space-like slices, which were referred to as locative slices in [41]. Such slices are not so large as to result in acausal influences, but large enough to capture nonlocal correlations. It was an open problem whether such slices can be captured by a deductive system, such that proof search corresponds to quantum evolution. In a joint work with Blute, Guglielmi, Ivanov, and Panangaden, Straßburger has shown that the logic BV with its mix of commutative and noncommutative connectives, is precisely the right logic for such analysis. More precisely, it was shown that the commutative tensor encodes (possible) entanglement, and the noncommutative *seq* encodes causal precedence. With this interpretation, the locative slices are precisely the derivable strings of formulas. Several new technical results about BV are developed as part of this analysis, which is published in [28]

6.7. On the Pigeonhole and Related Principles in Deep Inference and Monotone Systems

Participant: Anupam Das.

The size of proofs of the propositional pigeonhole principle over various systems is a topic of much interest in the proof complexity literature. In particular, it has received notable attention in recent years from the deep inference community, where its classification over the system KS appears as an open problem in numerous publications. In [21] we construct quasipolynomial-size proofs of the propositional pigeonhole principle in the deep inference system KS, addressing this question by matching the best known upper bound for the more general class of monotone proofs.

We make significant use of monotone formulae computing boolean threshold functions, an idea previously considered in works of Atserias et al. The main construction, monotone proofs witnessing the symmetry of such functions, involves an implementation of merge-sort in the design of proofs in order to tame the structural behavior of atoms, and so the complexity of normalization. Proof transformations from previous work on atomic flows are then employed to yield appropriate KS proofs.

As further results we show that our constructions can be applied to provide quasipolynomial-size KS proofs of the parity principle and the generalized pigeonhole principle. These bounds are inherited for the class of monotone proofs, and we are further able to construct $nO(\log\log n)$-size monotone proofs of the weak pigeonhole principle, thereby also improving the best known bounds for monotone proofs.

6.8. A multi-focused proof system isomorphic to expansion proofs

Participants: Kaustuv Chaudhuri, Stefan Hetzl [Vienna University of Technology, Vienna, Austria], Dale Miller.

The sequent calculus is often criticized for requiring proofs to contain large amounts of low-level syntactic details that can obscure the essence of a given proof. Because each inference rule introduces only a single connective, sequent proofs can separate closely related steps—such as instantiating a block of quantifiers—by irrelevant noise. Moreover, the sequential nature of sequent proofs forces proof steps that are syntactically non-interfering and permutable to nevertheless be written in some arbitrary order. The sequent calculus thus lacks a notion of *canonicity*: proofs that should be considered essentially the same may not have a common syntactic form. To fix this problem, many researchers have proposed replacing the sequent calculus with proof structures that are more parallel or geometric. Proof-nets, matings, and atomic flows are examples of such *revolutionary* formalisms. In [13], we propose, instead, an *evolutionary* approach to recover canonicity within the sequent calculus, which we illustrate for classical first-order logic. The essential element of our approach is the use of a *multi-focused* sequent calculus as the means for abstracting away low-level details from classical cut-free sequent proofs. We show that, among the multi-focused proofs, the *maximally multi-focused* proofs that collect together all possible parallel foci are canonical. Moreover, if we start with a certain focused sequent proof system, such proofs are isomorphic to *expansion proofs*—a well known, minimalistic, and parallel generalization of Herbrand disjunctions—for classical first-order logic. This technique appears to be a systematic way to recover the "essence of proof" from within sequent calculus proofs.

6.9. Equality and fixpoints in the calculus of structures

Participants: Kaustuv Chaudhuri, Nicolas Guenot [IT University of Copenhagen, Denmark].

The standard proof theory for logics with equality and fixpoints suffers from limitations of the sequent calculus, where reasoning is separated from computational tasks such as unification or rewriting. We propose in [20] an extension of the calculus of structures, a deep inference formalism, that supports incremental and contextual reasoning with equality and fixpoints in the setting of linear logic. This system allows deductive and computational steps to mix freely in a continuum which integrates smoothly into the usual versatile rules of multiplicative-additive linear logic in deep inference.

6.10. Automatically deriving schematic theorems for dynamic contexts

Participants: Kaustuv Chaudhuri, Olivier Savary-Bélanger [Princeton University, USA].

Hypothetical judgments go hand-in-hand with higher-order abstract syntax for meta-theoretic reasoning. Such judgments have two kinds of assumptions: those that are statically known from the specification, and the *dynamic assumptions* that result from building derivations out of the specification clauses. These dynamic assumptions often have a simple regular structure of repetitions of *blocks* of related assumptions, with each block generally involving one or several variables and their properties, that are added to the context in a single backchaining step. Reflecting on this regular structure can let us derive a number of structural properties about the elements of the context.

In [26], we present an extension of the Abella theorem prover, which is based on a simply typed intuitionistic reasoning logic supporting (co-)inductive definitions and generic quantification. Dynamic contexts are represented in Abella using lists of formulas for the assumptions and quantifier nesting for the variables, together with an inductively defined *context relation* that specifies their structure. We add a new mechanism for defining particular kinds of regular context relations, called *schemas*, and *tacticals* to derive theorems from these schemas as needed. Importantly, our extension leaves the trusted kernel of Abella unchanged. We show that these tacticals can eliminate many commonly encountered kinds of administrative lemmas that would otherwise have to be proven manually, which is a common source of complaints from Abella users.

6.11. A two-level logic approach for reasoning about typed specification languages

Participants: Kaustuv Chaudhuri, Mary Southern [University of Minnesota, USA].

The *two-level logic approach* (*2LLA*) to reasoning about computational specifications, as implemented by the Abella theorem prover, represents derivations of a *specification language* as an inductive definition in a *reasoning logic*. This approach has traditionally been formulated with the specification and reasoning logics having the *same* type system, and only the formulas being translated. However, requiring identical type systems limits the approach in two important ways: (1) every change in the specification language's type system requires a corresponding change in that of the reasoning logic, and (2) the same reasoning logic cannot be used with two specification languages at once if they have incompatible type systems. In [27], we propose a technique based on *adequate* encodings of the types and judgments of a typed specification language in terms of a simply typed higher-order logic program, which is then used for reasoning about the specification language in the usual *2LLA*. Moreover, a single specification logic implementation can be used as a basis for a number of other specification languages just by varying the encoding. We illustrate our technique with an implementation of the LF dependent type theory as a new specification language for Abella, co-existing with its current simply typed higher-order hereditary Harrop specification logic, without modifying the type system of its reasoning logic.

6.12. Undecidability of multiplicative subexponential logic

Participant: Kaustuv Chaudhuri.

Subexponential logic is a variant of linear logic with a family of exponential connectives—called *subexponentials*—that are indexed and arranged in a pre-order. Each subexponential has or lacks associated structural properties of weakening and contraction. In [18], we show that classical propositional multiplicative linear logic extended with one unrestricted and two incomparable linear subexponentials can encode the halting problem for two register Minsky machines, and is hence undecidable.

6.13. Meta-theoretic results on type isomorphisms in the presence of sums

Participant: Danko Ilik.

Type isomorphisms are a pervasive notion of Theoretical Computer Science. In functional programming, two data types being isomorphic means that we can coerce data and programs back-and-forth between two specifications without loss of information. In Constructive Mathematics, two sets are of the same cardinality exactly when they are isomorphic as types. In the proof theory of intuitionistic logic, two formulas are strongly equivalent precisely when they are isomorphic as types.

However, the theory of simple types made from functions, products, and sums, is well understood only when we do not treat functions and sums at the same time. Fiore, Di Cosmo, and Balat [50], presented a "negative" results: the theory of those type isomorphisms is not finitely axiomatizable. To establish the result, they used the work around the Tarski High School Algebra Problem from Mathematical Logic.

We showed that the picture is not so dark by presenting a positive result: the theory is recursively axiomatizable and decidable. The proofs exploit further the deep theory around Tarski's Problem. This work was presented at the Joint Meeting of the Twenty-Third EACSL Annual Conference on Computer Science Logic (CSL) and the Twenty-Ninth Annual ACM/IEEE Symposium on Logic in Computer Science (LICS) in Vienna, Austria [23].

6.14. Towards proof canonicity in presence of disjunction and induction

Participants: Hichem Chihani, Danko Ilik.

The previous work on type isomorphisms showed a way to treat the problem of identity/canonicity of proofs for intuitionistic logic with disjunction, or, equivalently, the problem of the (non-)existence of a canonical eta-long normal form for lambda calculus with if-expressions, which is a long standing open question.

One can see this from the perspective of focusing sequent calculi. The asynchronous phase of proof search is an oriented application of type isomorphisms (by the formulas-as-types correspondence). As we already know that, in the absence of disjunction (sum types), a cut-free focused derivation is eta-long and unique (when the data provided by the synchronous phase is the same), what is necessary in order to handle disjunction is to propagate isomorphisms further than what usual sequent calculus allows. This is related in spirit to deep inference, but more conservative. An implementation of a canonical normalizer and a paper on the topic is under way.

We also intend to use the method to give a proof of focused cut-elimination for the sequent calculi LJF and LKF (at least, for the Sigma-2 fragment) extended with induction. A formal proof in Agda is under development.

6.15. Interpretation of the Sigma-2-classical Axiom of Choice in System T

Participant: Danko Ilik.

Updating previous work, we showed that one can develop a realizability interpretation for the Σ_2^0-fragment of classical Analysis in System T only [36].

This is known to be possible, in principle, by a 1979 result of Schwichtenberg. However, up to day no method that avoids both bar recursion (Spector) and control operators (Krivine) has been known. In fact, we propose to treat control operators as a meta-mathematical technique, rather than to have them in the language of realizers as classical realizability does; we provide a formal proof in Agda that control operators can be completely normalized away from System T while preserving essential equations. [15]

6.16. Axiomatization of constraint systems for first-order reasoning modulo a theory

Participants: Damien Rouhling, Stéphane Graham-Lengrand, Assia Mahboubi, Jean-Marc Notin, Mahfuza Farooque.

This result is part of a work in theorem proving, whose purpose is to provide a theoretical basis for the handling of quantifiers in presence of a theory for which we have specific decision procedures. Inspired by the way first-order unifiers are generated and propagated in automated reasoning techniques such as *tableaux* methods, we sought to generalise these mechanisms to the presence of a theory: We introduced a axiomatic notion of constraint system and a sequent calculus introducing meta-variables and propagating constraints. We then identified the axioms that should be satisfied by the theory's decision procedure, in order for the sequent calculus to be sound and complete. This provides the theoretical basis for the development of Psyche 2.0. This result is submitted for publication.

6.17. Realisability models for cut-elimination in focused systems

Participant: Stéphane Graham-Lengrand.

This result is part of the effort to build meaningful semantics for classical proofs, here based on a polarisation of logical formulae: positive or negative.

Following work by Zeilberger [80], a computational interpretation of cut-elimination in the focused systems LJF and LKF can be given: proofs of positive formulae provide structured data, while proofs of negative formulae consume such data; focusing allows the description of the interaction between the two kinds of proofs as pure pattern-matching.

First, we showed this at a level of abstraction where formulae are no longer made of syntax, yet we also extended the approach so that it could treat quantifiers.

Second, we connected this interpretation to realisability semantics, more precisely orthogonality models, where positive formulae are interpreted as sets of data, and negative formulae are interpreted as their orthogonal sets.

Our construction of orthogonality models for the focused systems LKF and LJF describe the pattern-matching process of cut-elimination in terms of orthogonality. This result has been proved in the Coq proof assistant and forms the second part of [11].

6.18. Refining the FPC framework

Participants: Roberto Blanco, Zakaria Chihani, Quentin Heath, Dale Miller, Fabien Renaud.

We have continued to develop our approach to Foundational Proof Certificates (FPCs). This framework allows defining proof evidence in a general fashion. Proofs in both intuitionistic and classical logics are definable in this framework. We originally have written two different kernels for checking these results but more recently we have found that we can exploit an encoding due to Chaudhuri [43] that enables us to only implement the intuitionistic kernel and then simply encode the classical formulas so that they operator directly on the intuitionistic kernel. This encoding allows for a much more precise and simple means for encoding classical logic into intuitionistic logic than the more familiar double negation translations.

We have also started to develop the second phase of defining proof evidence that was proposed in the ProofCert proposal: the definition of proofs that require fixed points (induction / co-induction). We now have two different kernels being developed on top of the Bedwyr model checker that are checking (and in some cases, proving) theorems involving induction, reachability, and bisimulation.

6.19. Structuring a refinement engine using logic programming

Participants: Dale Miller, Claudio Sacerdoti Coen [University of Bologna], Enrico Tassi [MSR Inria Joint Lab].

The Matita theorem prover is an implementation of the Calculus of Inductive Constructions that is meant to be more accessible (as an implementation) than the Coq system. In an effort to make the Matita kernel more accessible and more flexible, the implementers of that system are experimenting with using a logic programming language similar to λProlog as the control system of the refinement mechanism. In order to use such a logic programming language in this capacity, the notion of flexible goal suspension and *when* declarations are needed. Such a λProlog re-implementation has been written and some experiments in deploying such a system are underway. Formal aspects of λProlog specifications have also been performed using the Abella theorem prover.

PERCEPTION Project-Team

5. New Results

5.1. Highlights of the Year

- In 2014 Antoine Deleforge (team member 2009-2013) received the **Signal, Image and Vision best PhD prize** for his thesis *"Acoustic Space Mapping: A Machine Learning Approach to Sound Source Separation and Localization"*, defended in December 2013 and advised by Radu Horaud. The prize is jointly awarded by GDR ISIS, Club EEA, and GRETSI.

  Website: http://www.inria.fr/centre/grenoble/actualites/apprendre-a-rester-attentif-a-ses-locuteurs
- Radu Horaud was awarded an **ERC Advanced Grant** for his five year project VHIA *"Vision and Hearing in Action"*, grant number 340113, 2014-2019.

  Website: https://team.inria.fr/perception/vhia/.
- The PERCEPTION team was awarded an **ANR BLANC** two year project MIXCAM *"Real-Time Visual Reconstruction by Mixing Multiple Depth and Color Cameras"*, in collaboration with 4D View Solutions, 2014-2016.

  Website: https://team.inria.fr/perception/mixcam-project/
- The PERCEPTION team was awarded an **FP7 STREP** three year project EARS *"Embodied Audition for Robots"*, in collaboration with Friedrich Alexander Universiteit, coordinator (Germany), Ben Gurion University (Israel), Imperial College (UK), Humboldt University Berlin (Germany) and Aldebaran Robotics (France), 2014-2017.

  Website: https://team.inria.fr/perception/ears/

5.2. Acoustic Space Learning on Binaural Manifolds

We addressed the problems of modeling the acoustic space generated by a full-spectrum sound source and of using the learned model for the localization and separation of multiple sources that simultaneously emit sparse-spectrum sounds. We lay theoretical and methodological grounds in order to introduce the binaural manifold paradigm. We perform an in-depth study of the latent low-dimensional structure of the high-dimensional interaural spectral data, based on a corpus recorded with a human-like audiomotor robot head, namely the POPEYE robot shown on Fig 3 (right). A non-linear dimensionality reduction technique is used to show that these data lie on a two-dimensional (2D) smooth manifold parameterized by the motor states of the listener, or equivalently, the sound source directions, e.g., Fig. 4 . We propose a probabilistic piecewise affine mapping model (PPAM) specifically designed to deal with high-dimensional data exhibiting an intrinsic piecewise linear structure. We derive a closed-form expectation-maximization (EM) procedure for estimating the model parameters, followed by Bayes inversion for obtaining the full posterior density function of a sound source direction. We extend this solution to deal with missing data and redundancy in real world spectrograms, and hence for 2D localization of natural sound sources such as speech. We further generalize the model to the challenging case of multiple sound sources and we propose a variational EM framework. The associated algorithm, referred to as variational EM for source separation and localization (VESSL) yields a Bayesian estimation of the 2D locations and time-frequency masks of all the sources. Comparisons of the proposed approach with several existing methods reveal that the combination of acoustic-space learning with Bayesian inference enables our method to outperform state-of-the-art methods [18], [24].

Website: https://team.inria.fr/perception/research/acoustic-learning/

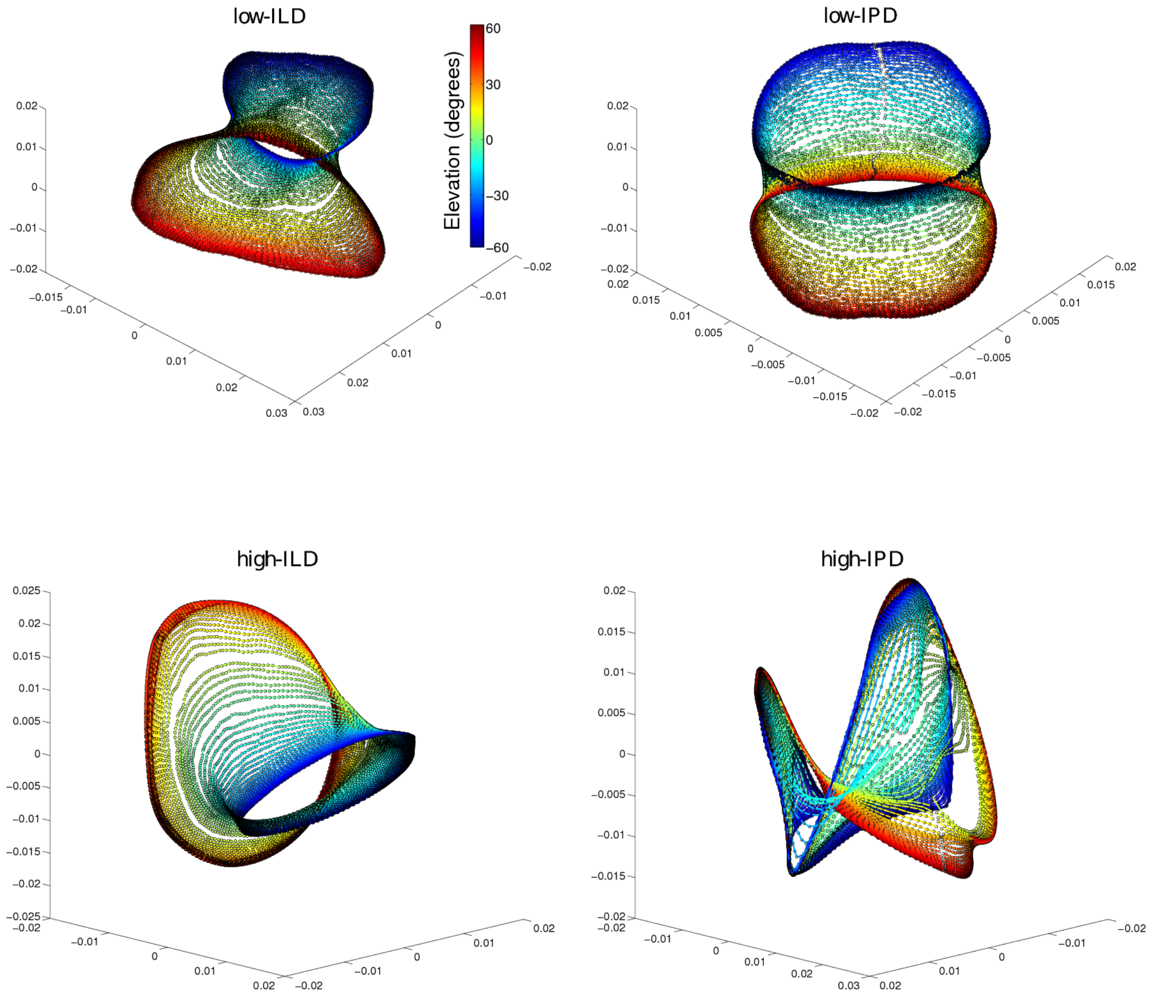


Figure 4. This figure illustrates the concept of binaural manifold. A wide-spectrum sound is recorded with a binaural acoustic dummy head and an interaural high-dimensional spectral representation of this sound is mapped onto a low-dimensional (2) space. This reveals the two-dimensional manifold of possible sound-source directions that is embedded in the interaural spectral features. Please consult [18] for more details.

5.3. Geometric Sound Source Localization

We address the problem of sound-source localization from time-delay estimates using arbitrarily-shaped non-coplanar microphone arrays. A novel geometric formulation is proposed, together with a thorough algebraic analysis and a global optimization solver [15]. The proposed model is thoroughly described and evaluated. The geometric analysis, stemming from the direct acoustic propagation model, leads to necessary and sufficient conditions for a set of time delays to correspond to a unique position in the source space. Such sets of time delays are referred to as *feasible sets*. We formally prove that every feasible set corresponds to exactly one position in the source space, whose value can be recovered using a closed-form localization mapping. Therefore we seek for the optimal feasible set of time delays given, as input, the received microphone signals. This time delay estimation problem is naturally cast into a programming task, constrained by the feasibility conditions derived from the geometric analysis. A global branch-and-bound optimization technique is proposed to solve the problem at hand, hence estimating the best set of feasible time delays and, subsequently, localizing the sound source. Extensive experiments with both simulated and real data are reported; we compare our methodology to four state-of-the-art techniques. This comparison shows that the proposed method combined with the branch-and-bound algorithm outperforms existing methods. These in-depth geometric understanding, practical algorithms, and encouraging results, open several opportunities for future work.

Website: https://team.inria.fr/perception/research/geometric-sound-source-localization/

5.4. Joint Registration of Multiple Point Sets

We developed a probabilistic generative model and its associated algorithm to jointly register multiple point sets. The vast majority of state-of-the-art registration techniques select one of the sets as the *model* and perform pairwise alignments between the other sets and this set. The main drawback of this mode of operation is that there is no guarantee that the model-set is free of noise and outliers, which contaminates the estimation of the registration parameters. Unlike previous work, the proposed method treats all the point sets on an equal footing: they are realizations of a Gaussian mixture (GMM) and the registration is cast into a clustering problem [26]. We formally derive an EM algorithm that estimates both the GMM parameters and the rotations and translations that map each individual set onto the *central* model. The mixture means play the role of the registered set of points while the variances provide rich information about the quality of the registration. We thoroughly validate the proposed method with challenging datasets, we compare it with several state-of-the-art methods, and we show its potential for fusing real depth data.

Website: https://team.inria.fr/perception/research/jrmpc/

5.5. High-Dimensional Regression

The problem of approximating high-dimensional data with a low-dimensional representation is addressed. The article makes the following contributions. An inverse regression framework is proposed, which exchanges the roles of input and response, such that the low-dimensional variable becomes the regressor, and which is tractable. A mixture of locally-linear probabilistic mapping model is introduced, that starts with estimating the parameters of the inverse regression, and follows with inferring closed-form solutions for the forward parameters of the high-dimensional regression problem of interest. Moreover, a partially-latent paradigm is introduced, such that the vector-valued response variable is composed of both observed and latent entries, thus being able to deal with data contaminated by experimental artifacts that cannot be explained with noise models. The proposed probabilistic formulation could be viewed as a latent-variable augmentation of regression. Expectation-maximization (EM) procedures are introduced, based on a data augmentation strategy which facilitates the maximum-likelihood search over the model parameters. Two augmentation schemes are proposed and the associated EM inference procedures are described in detail; they may well be viewed as generalizations of a number of EM regression, dimension reduction, and factor analysis algorithms. The proposed framework is validated with both synthetic and real data. Experimental evidence is provided that the method outperforms several existing regression techniques [19], [25].

Website: https://team.inria.fr/perception/research/high-dim-regression/

5.6. Audiovisual Speaker Detection, Localization and Interaction with NAO

In this research we address the problem of audio-visual speaker detection. We introduce an online system working on the humanoid robot NAO. The scene is perceived with two cameras and two microphones. A *multimodal* Gaussian mixture model (GMM) fuses the information extracted from the auditory and visual sensors. The system is implemented based on a platform-independent middleware library and it is able to process the information online (17 visual frames per second). A detailed method description and the system implementation are provided, with special emphasis on the online processing issues that must be addressed, and the proposed solutions. Experimental validation is done over five different scenarios, with no special lighting, nor special acoustic conditions, leading to good results [16].

Website: https://team.inria.fr/perception/research/audiovisual-nao/

5.7. EM for Weighted-Data Clustering

Figure 5. We developed a novel multimodal clustering method that is based on expectation-maximization (EM) with weighted data. The left image shows auditory features (green), namely sound source positions mapped onto the image plane using [24] and visual features (blue, lip landmarks), as well as the active speaker (yellow square). The right image shows the results of our weighted-data EM algorithm that finds three clusters. Among these clusters, the active audio-visual cluster is marked with a transparent blue circle.

Data clustering has received a lot of attention and many methods, algorithms and software packages are currently available. Among these techniques, parametric finite-mixture models play a central role due to their interesting mathematical properties and to the existence of maximum-likelihood estimators based on expectation-maximization (EM). In this paper we propose a new mixture model that associates a weight with each observed data point. We introduce a Gaussian mixture with weighted data and we derive two EM algorithms [29]: the first one considers the weight of each observed datum to be fixed, while the second one treats each weight as a hidden variable drawn from a gamma distribution. We provide a general-purpose scheme for weight initialization and we thoroughly validate the proposed algorithms by comparing them with several parametric and non-parametric clustering techniques. We demonstrate the utility of our method for clustering heterogeneous data, namely data gathered with different sensorial modalities, e.g., audio and vision.

Website: https://team.inria.fr/perception/research/wdgmm/

5.8. Continuous Action Recognition

Continuous action recognition is more challenging than isolated recognition because classification and segmentation must be simultaneously carried out. We build on the well known dynamic time warping (DTW) framework and devise a novel visual alignment technique, namely dynamic frame warping (DFW), which performs isolated recognition based on per-frame representation of videos, and on aligning a test sequence with a model sequence. Moreover, we propose two extensions which enable to perform recognition concomitant with segmentation, namely one-pass DFW and two-pass DFW. These two methods have their roots in the domain of continuous recognition of speech and, to the best of our knowledge, their extension to continuous visual action recognition has been overlooked. We test and illustrate the proposed techniques with a recently released dataset (RAVEL) [32] and with two public-domain datasets widely used in action recognition (Hollywood-1 and Hollywood-2). We also compare the performances of the proposed isolated and continuous recognition algorithms with several recently published methods [22].

Website: https://team.inria.fr/perception/research/car/

5.9. Skeletal Quads

Recent advances on human motion analysis have made the extraction of human skeleton structure feasible, even from single depth images. This structure has been proven quite informative for discriminating actions in a recognition scenario. In this context, we propose a local skeleton descriptor that encodes the relative position of joint quadruples. Such a coding implies a similarity normalization transform that leads to a compact (6D or 5D) view-invariant skeletal feature, referred to as skeletal quad. In the references below, we use this descriptor in conjunction with Fisher kernel in order to encode gesture or action (sub)sequences. The short length of the descriptor compensates for the large inherent dimensionality associated to Fisher vectors. We investigate the performance in both isolated [28] and continuous [27] recognition scenarios.

Website: https://team.inria.fr/perception/research/skeletalquads/

PHOENIX Project-Team

6. New Results

6.1. Highlights of the Year

- A best paper award was obtained at ASSETS 2014 (The 16th International ACM SIGACCESS Conference on Computers and Accessibility), by the 5 authors of the paper "Tablet-Based Activity Schedule for Children with Autism in Mainstream Environment" .

BEST PAPERS AWARDS :

[26] **ASSETS 2014 - The 16th International ACM SIGACCESS Conference on Computers and Accessibility**. C. FAGE, L. POMMEREAU, C. CONSEL, E. BALLAND, H. SAUZÉON.

6.2. Technological Support for Self-Regulation of Children with Autism

Children with Autism Spectrum Disorders (ASD) have difficulties to self-regulate emotions, impeding their inclusion in a range of mainstreamed environments. Self-regulating emotions has been shown to require recognizing emotions and invoking specific coping strategies.

In the context of the School+ research project, we have developed an application dedicated to self-regulating emotions in children with ASD. Ten children with ASD have experimentally tested this tablet-based application over a period of three months in a mainstreamed school. A collaborative learning approach, involving parents, teachers and a school aid, was used 1) to train students to operate the tablet and our application autonomously, and 2) to facilitate the adoption of our intervention tool.

This study shows that our application was successful in enabling students with ASD to self-regulate their emotions in a school environment. Our application helped children with autism to recognize and name their emotions, and to regulate them using idiosyncratic, parent-child, coping strategies, supported by multimedia contents.

This work is in the context of the School+ national research project funded by the French Ministry of National Education.

A best paper award was obtained at ASSETS 2014 (The 16th International ACM SIGACCESS Conference on Computers and Accessibility) for this work in October 2014, by the 5 authors of the paper "Tablet-Based Activity Schedule for Children with Autism in Mainstream Environment" [26]: Charles Fage, Léonard Pommereau, Charles Consel, Emilie Balland, and Hélène Sauzéon.

6.3. A Low-Cost approach to the Verification of Daily Activities of Elders

Activities of Daily Living (ADL) are abilities defining the functional status of an individual. Verifying what ADLs are performed by an elder is a decisive factor to determine what kinds and what levels of assistance are needed for an individual and whether aging in place is desirable. The importance of this issue has led a number of researchers to develop a range of Ubicomp approaches that can monitor activities.

In this study, we take these prior results one step further and apply them to the needs of caregiver professionals to monitor elders at their home. Specifically, our approach relies on the following key observation: as people age their daily activities are increasingly organized according to a routine to optimize their daily functioning. As a result, their activities do not need to be recognized but should rather be verified. Deviations are a warning sign of degradation.

We have developed an approach to activity verification. This approach relies on a technological infrastructure that is simple, low-cost and non-intrusive. This infrastructure was deployed in four homes of elders of 83 years of age on average. The same set of sensors was used in the four homes and was placed at strategic locations with respect to their routines to verify the target activities. The analysis of the data collected during five weekdays show that they follow very strict routines that can easily be associated with their main activities.

This work is in the context of the DomAssist project, funded by the following partners: UDCCAS, CG33, CRA, CNSA, Chambre des métiers. A report of the work has been published at the 16th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS 2014) [25].

6.4. Using virtual reality for studying everyday-like memory and its cognitive correlates

This work consisted in a pilot-study with a comparison approach between aging and traumatic brain injury (TBI) to investigate everyday object memory patterns using a virtual HOMES test.

- Methods: Sixteen young controls, 15 older adults and 15 TBI patients underwent the HOMES test and traditional tests.
- Results: Older adults and TBI patients exhibited similar HOMES performances: poor recall, a greater recognition benefit, high false recognitions, but intact clustering and proactive interference effects. The age-related differences for HOMES measures were mainly mediated by executive functioning, while the HOMES performances in the TBI group were correlated with memory measures.
- Conclusion: The differential cognitive mediating effects for a similar everyday-like memory pattern have been discussed by highlighting the need for more cautious interpretations of cognitive mechanisms behind similar behavioral patterns in different populations especially in clinical and rehabilitation settings.
- Implication for Rehabilitation:
  - Virtual reality might provide ecological scenarios to assess the multiple processes of everyday memory in elderly people as well as in TBI patients.
  - A similar pattern of Everyday-like memory failures might result from different cognitive origins among different neuropsychological patients.
  - The assessment of specific cognitive origins of Everyday-like memory impairments deserves consideration for drawing up relevant rehabilitative programs that match the specific cognitive needs of patients for performing everyday memory tasks.

This work has been published in the journal "Disability and Rehabilitation: Assistive Technology" in November 2014 [14].

PI.R2 Project-Team

5. New Results

5.1. Highlights of the Year

We successfully organised the thematic trimester Semantics of Proofs and Certified Mathematics (IHP, April-July 2014). The trimester attracted over two hundred participants altogether (with about 60 “resident” participants staying a month or more), hosted 5 special workshops, as well as other related regevents such as Types, MAP (Mathematics, Algorithms, and Proofs). It was the first thematic trimester in the history of IHP to feature computer science prominently. There was a kick-off day on April 22, with talks of Georges Gonthier, Thomas Hales, Xavier Leroy, and Vladimir Voevodsky, with the presence of some science journalists. During the trimester, the Bourbaki Seminar devoted an afternoon (June 21) to these themes, with talks of Thomas Hales and Thierry Coquand.

Shortly before, Coq has received the Software System Award 2013 from the Association for Computing Machinery (ACM). Hugo Herbelin is one of the recipients of this prize.

5.2. Proof-theoretical and effectful investigations

Participants: Pierre Boutillier, Guillaume Claret, Pierre-Louis Curien, Amina Doumane, Hugo Herbelin, Etienne Miquey, Ludovic Patey, Pierre-Marie Pédrot, Yann Régis-Gianas, Alexis Saurin.

5.2.1. Proving with side-effects

In 2012, Hugo Herbelin showed that classical arithmetic in finite types extended with strong elimination of existential quantification proves the axiom of dependent choice. To get classical logic and choice together without being inconsistent is made possible first by constraining strong elimination of existential quantification to proofs that are essentially intuitionistic and secondly by turning countable universal quantification into an infinite conjunction of classical proofs evaluated along a call-by-need evaluation strategy so as to extract from them intuitionistic contents that complies to the intuitionistic constraint put on strong elimination of existential quantification. Étienne Miquey is currently working to get a presentation of this work in Curien-Herbelin's μ-$\widetilde{\mu}$-calculus, with the aim of getting in the end a CPS-translation. Such a translation would provide a strong argument of normalisation for the calculus, as well as a better undertanding of the mechanisms of the calculus, especially the side-effect part and the meaning of the existential quantifier restriction.

Hugo Herbelin and Danko Ilik carried on their work on the computational content of completeness proofs and in particular of the computational content of Gödel's completeness theorem. Hugo Herbelin presented their work at the workshop PSC 2014.

5.2.2. Reverse mathematics

Ludovic Patey studied with Laurent Bienvenu and Paul Shafer the provability strength of Ramsey-type versions of theorems like König's lemma. The corresponding paper is submitted to the Journal of Mathematical Logic. Ludovic Patey studied with Laurent Bienvenu the constructions of diagonal non-computable functions by probabilistic means. They submitted a paper to Information and Computation. Ludovic Patey worked on the existence of universal instances in reverse mathematics, and submitted a paper to Annals of Pure and Applied Logic. He worked on the relations between diagonal non-computability and Ramsey-type theorems and submitted a paper to the Archive for Mathematical Logic. He studied the links between the iterative forcing framework developed by Lerman, Solomon & Towsner and the notion of preservation of hyperimmunity and submitted a paper to Computability in Europe 2015.

5.2.3. Gödel's functional interpretation

Pierre-Marie Pédrot kept developing the proof-as-program interpretation of Gödel's Dialectica translation, as seen through the prism of classical realisability. This work was presented at TYPES 2014 and later published at LICS 2014 [26].

5.2.4. Logical foundations of call-by-need evaluation

Alexis Saurin and Pierre-Marie Pédrot developed a structured reconstruction of call-by-need based on linear head reduction which arose in the context of linear logic. This opens new directions both to extend call-by-need to control and to apply linear logic proof-theory (and particularly proof-nets) to call-by-need evaluation. This work was presented at JFLA 2014 [30] early 2014 and later expanded to the classical case, encompassing $\lambda\mu$-calculus.

5.2.5. Streams and classical logic

Alexis Saurin and Fanny He have been working on transfinite term rewriting in order to model stream calculi and their connections with lambda-calculi for classical logic. Their work gave rise to a presentation at the Workshop on Infinitary Rewriting that took place in Vienna last July as part of FLOC 2014.

5.2.6. Alternative syntaxes for proofs

Amina Doumane and Alexis Saurin, in a joint work with Marc Bagnol, studied the structure of several correctness criteria for linear logic proof-nets and could relate them through a new primitive notion of dependency. This work was first presented at JFLA 2014 [29] early 2014 and later at Structure and Deduction in Vienna as part of FLOC 2014. An expanded version has recently been accepted at FOSSACS 2015 [19].

5.3. Type theory and the foundations of Coq

Participants: Pierre Boutillier, Pierre-Louis Curien, Hugo Herbelin, Pierre-Marie Pédrot, Yann Régis-Gianas, Matthieu Sozeau, Arnaud Spiwack.

5.3.1. Description of type theory

Hugo Herbelin and Arnaud Spiwack completed and published their characterisation of the type constructions of Coq in terms of atomic constructions rather than their usual description as a monolithic scheme [23]. This work permitted both a more pedagogical presentation of Coq's type system, and a more tractable and composable mathematical model of Coq on which meta-properties can be stated and proved.

5.3.2. Models of type theory

Simplicial sets and their extensions as Kan complexes can serve as models of homotopy type theory. Hugo Herbelin developed a concrete type-theoretic formalisation of semi-simplicial sets following ideas from Steve Awodey, Peter LeFanu Lumsdaine and other researchers both at Carnegie-Mellon University and at the Institute of Advanced Study. This is in the process of being published in a special issue of MSCS on homotopy type theory [9].

The technique scales to provide type-theoretic constructions for arbitrary presheaves on Reedy categories, thus including simplicial sets.

5.3.3. Proof irrelevance, eta-rules

During his master's internship supervised by Matthieu Sozeau, Philipp Haselwarter studied a formulation of proof-irrelevance based on the rooster and the syntactic bracket presentation by Spiwack and Herbelin [23]. This resulted in a decomposition of the calculus cleanly showing the use of smashing and a better understanding of the restricted elimination rules of propositions. It also clearly shows that the inductive type for accessibility, used to justify general wellfounded definitions, can not be interpreted as a proof-irrelevant proposition in this calculus.

5.3.4. Unification

Matthieu Sozeau is continuing work in collaboration with Beta Ziliani (PhD at MPI-Saarbrücken) on formalising the unification algorithm used in Coq, which is central for working with advanced type inference features like Canonical Structures. This is the first precise formalisation of all the rules of unification including the ones used for canonical structure resolution. The presentation currently excludes some heuristics that were added on top of the core algorithm in Coq, until they can be studied more carefully. This work, part of B. Ziliani's thesis, was presented at the UNIF'14 workshop [28] and the Coq workshop in Vienna. A submission is in preparation.

5.3.5. Foundations and paradoxes

Arnaud Spiwack generalised previous works by Herman Geuvers and Hugo Herbelin to implement Hurkens's paradox of the impredicative system U^-. The resulting Coq implementation, which is completely independent from the impredicative features of Coq, generalises the two special cases which were previously used to prove negative results about impredicativity in Coq.

5.4. Homotopy of rewriting systems

Participants: Cyrille Chenavier, Pierre-Louis Curien, Yves Guiraud, Maxime Lucas, Philippe Malbos, Jovana Obradović.

5.4.1. Coherent presentations of Artin monoids

With Stéphane Gaussent (ICJ, Univ. de Saint-Étienne), Yves Guiraud and Philippe Malbos have used higher-dimensional rewriting methods for the study of Artin monoids, a class of monoids that is fundamental in algebra and geometry. This work uses the formal setting of coherent presentations (a truncation of polygraphic resolutions at the level above relations) to formulate, in a common language, several known results in combinatorial group theory: one by Tits about the fundamental group of a graph associated to an Artin monoid [65], and one by Deligne about the actions of Artin monoids on categories [47], both proved by geometrical methods. In this work, an improvement of Knuth-Bendix's completion procedure is introduced, called the homotopical completion-reduction procedure, and it is used to give a constructive proof and to extend both theorems. This work will appear in Compositio Mathematica [18] and has been implemented in a Python library.

The next objective of this collaboration is to extend those results in every dimension, first to Artin monoids, then to Artin groups, with a view towards two well-known open problems in the field: the word problem of Artin groups and the so-called $K(\pi, 1)$ conjecture.

5.4.2. New methods for the computation of polygraphic resolutions

Maxime Lucas, supervised by Pierre-Louis Curien and Yves Guiraud, develops Squier's theory in the setting of cubical ω-categories. This will allow easier and more explicit computations of polygraphic resolutions than in the globular setting of [5], and the use of new effective methods such as the reversing algorithm from Garside theory [46].

Yves Guiraud currently collaborates with Patrick Dehornoy (Univ. de Caen) and Matthieu Picantin (LIAFA, Univ. Paris 7) to extend the constructions of [18] to other important families of monoids, such as the plactic monoid, the Chinese monoid and the dual braid monoids.

5.4.3. Higher-dimensional linear rewriting

Cyrille Chenavier, Pierre-Louis Curien, Yves Guiraud and Philippe Malbos investigate with Eric Hoffbeck (LAGA, Univ. Paris 13) and Samuel Mimram (LIX, École Polytechnique) the links between set-theoretic rewriting theory and the computational methods known in symbolic algebra, such as Gröbner bases [39]. This interaction is supported by the Focal project of the IDEX Sorbonne Paris Cité.

With Eric Hoffbeck (LAGA, Univ. Paris 13), Yves Guiraud and Philippe Malbos have introduced the setting of linear polygraphs to formalise a theory of linear rewriting (in the sense of linear algebra), generalising Gröbner bases. They have adapted to algebras the procedure of [5] that computes polygraphic resolutions from convergent presentations of monoids, with applications to the decision of an important homological property called Koszulness. This work is contained in [35] and it has been presented at IWC 2014 [31].

Cyrille Chenavier, supervised by Yves Guiraud and Philippe Malbos, explores the use of Berger's theory of reduction operators [38] to design new methods for the study of linear rewriting systems, and to promote the use of rewriting techniques in combinatorial algebra.

5.4.4. *Homotopical and homological finiteness conditions*

Yves Guiraud and Philippe Malbos have written a comprehensive introduction [36] on the links between higher-dimensional rewriting, the homotopical finiteness condition "finite derivation type" and the homological finiteness condition "FP_3", from the point of view of higher categories and polygraphs. The purpose of this work is to provide an introduction to the field, formulated in a contemporary language, and with new, more formal proofs of classical results.

5.4.5. *Wiring structure of operads and operad-like structures*

Building on recent ideas of Marcelo Fiore on the one hand, and of François Lamarche on the other hand, Pierre-Louis Curien and Jovana Obradović develop a syntactic approach, using some of the kit of Curien-Herbelin's duality of computation and its polarised versions of Munch and Curien, to the definition of various structures that have appeared in algebra under the names of operads, cyclic operads, dioperads, properads, modular and wheeled operads, permutads, etc.... These structures are defined in the literature in different flavours. We seek to formalise the proofs of equivalence between these different styles of definition, and to make these proofs modular, so as not to repeat them for each variation of the notion of operad. Preliminary results are being presented in January 2015 at the Mathematical Institute of the Academy of Sciences (Belgrade).

5.5. Coq as a functional programming language

Participants: Pierre Boutillier, Guillaume Claret, Lourdes Del Carmen González Huesca, Thibaut Girka, Hugo Herbelin, Pierre Letouzey, Matthias Puech, Yann Régis-Gianas, Matthieu Sozeau, Arnaud Spiwack.

5.5.1. *Type classes and libraries*

Type Classes are heavily used in the HoTT/Coq library (http://github.com/HoTT/coq) started by the Univalent Foundations program at the IAS, to which Matthieu Sozeau participated. To ease the development of this sophisticated library, Matthieu Sozeau implemented a number of extensions to type class resolution to make it more predictable and efficient. These are now part of the Coq 8.5 release.

5.5.2. *Dependent pattern-matching*

The dissertation of Pierre Boutillier presents and formalises a new algorithm to compile dependent pattern-matching into a chain of Coq case analyses. It avoids the use of the "uniqueness of identity proofs" axiom in more cases than the former proposal by McBride and McKinna.

5.5.3. *Incrementality in proof languages*

Lourdes del Carmen González Huesca and Yann Régis-Gianas developed a new variant of the differential lambda calculus that has two main features: (i) it is deterministic ; (ii) it is based on a notion of a first-class changes. A paper is in preparation.

5.5.4. *Proofs of programs in Coq*

In collaboration with David Mentre (Mitsubishi Rennes), Thibaut Girka and Yann Régis-Gianas worked on a certified generator for correlating programs. A correlating program is a program that represents the semantic difference between two (close) versions of a program by performing a static scheduling of their instructions. Performing an abstract interpretation on the correlating program provides a representation of the semantic differences between the two versions of a program. A paper is written and should be submitted soon.

5.5.5. Typed tactic language

In collaboration with Beta Ziliani (MPI) and Thomas Refis (master 2 student at University Paris Diderot), Yann Régis-Gianas starts the development of the version 2 of Mtac, a tactic language for Coq. Mtac is a DSL embedded in the Coq proof assistant. Roughly speaking, it allows Coq to be used as a tactic language for itself. With this work, Mtac 2 now includes first class goals. A paper is in preparation.

5.5.6. Tactic engine

Arnaud Spiwack joined the team for two months (Sept—Oct 2014) to finalise the integration and documentation of his re-engineering of Coq's interactive proof engine for the v8.5 version. The new perspective taken by this new engine is to shift the primary focus from how tactics (proof instructions) can modify goals (proof obligations) to focus on the way tactics compose. By making sure that composition of tactics has good mathematical properties, the new engine makes it possible to combine tactics in a more predictable and more powerful way. This new engine is also notable for the introduction of an abstract interface for tactics and tactic composition which makes it easy to augment tactics with new capabilities. The most notable such features are so-called dependent subgoals, which makes more fine-grained proofs possible and significantly improves the support for dependent types; and backtracking which gives the possibility to deploy very modular proof-search components. During his two months in the team, Arnaud Spiwack also added support for tracing tactic execution (Info), again taking advantage of his modular design.

5.5.7. Effectful programming

Guillaume Claret and Yann Régis-Gianas developed a compiler from a subset of OCaml with effects to Coq. Possible effects are the exceptions, the global references and the non-termination. Guillaume Claret and Yann Régis-Gianas developed Pluto, a concurrent HTTP web server written in Gallina. They worked on techniques to certify such interactive programs, formalising the reasoning by use cases. Use cases are proven correct giving a scenario, a typed schema of interactions between a program and an environment, built using the tactic mode of Coq as a symbolic debugger.

5.5.8. Libraries

Sébastien Hinderer and Pierre Letouzey contributed an extended library of lists. Pierre Letouzey contributed an extended library about Peano numbers, that takes advantages of the "Numbers" modular framework done earlier.

POEMS Project-Team

6. New Results

6.1. Wave propagation in non classical media

6.1.1. Plasmonic black-hole waves at corners of metals

Participants: Anne-Sophie Bonnet-Ben Dhia, Camille Carvalho, Patrick Ciarlet.

This work, which is a part of the PhD of Camille Carvalho, is done in collaboration with Lucas Chesnel from CMAP at Ecole Polytechnique. We study the scattering of time-harmonic electromagnetic waves by a metallic obstacle in a 2D setting, at frequencies such that the dielectric permittivity of the metal has a negative real part and a small imaginary part. When the obstacle has corners, due to the sign-changing real part of the permittivity, unusual strong singularities for the electromagnetic field can appear. If the material dissipation is neglected, it can be necessary to consider a new functional framework, containing these singularities, to derive a well-posed problem. In this new framework, everything happens like if plasmonic waves were propagating to the corners, and a part of the energy is trapped by the corner, even if the material has been supposed non-dissipative. We have implemented an original numerical method consisting in using Perfectly Matched Layers at the corners to capture these black-hole waves. We have also proposed a new rule to mesh the corner in order to achieve convergence of classical finite elements in the simpler case where the problem is still well-posed in the classical framework. Finally, in collaboration with André Nicolet and Frédéric Zolla from Institut Fresnel in Marseille, we are now considering realistic dissipative metals. We show that there is still a significant effect of the black-hole phenomenon, which results in an unsual energy leakage in some frequency range.

6.1.2. Limiting amplitude principle for a two-layered dielectric/metamaterial medium

Participants: Maxence Cassier, Christophe Hazard, Patrick Joly.

This work has been a part of the PhD of Maxence Cassier and has allowed to initiate a collaboration with Boris Gralak from Institut Fresnel. For wave propagation phenomena, the limiting amplitude principle holds if the time-harmonic regime represents the large time asymptotic behavior of the solution of the evolution problem with a time-harmonic excitation. Considering a two-layered medium composed of a dielectric material and a Drude metamaterial separated by a plane interface, we prove that the limiting amplitude principle holds except for a critical situation related to a surface resonance phenomenon. Then the solution can either converge to the superposition of two time-periodic fields, or blow up linearly in time.

6.1.3. Perfectly Matched Layers in plasmas and metamaterials

Participants: Eliane Bécache, Patrick Joly, Maryna Kachanovska, Valentin Vinoles.

This work is a part of the PhD of Valentin Vinoles and is the subject of the post-doc of Maryna Kachanovska. It deals with the stability of Perfectly Matched Layers (PMLs) in dispersive media and is motivated by the fact that classical PMLs are unstable in negative index metamaterials and in some anisotropic plasmas. This led us to derive a new necessary criterion of stability which is valid for a large class of dispersive models and for more general PMLs than the classical ones. This criterion has been used to design new stable PMLs for negative index metamaterials and uniaxial anisotropic plasmas.

6.1.4. Retrieval method for anisotropic metamaterials

Participants: Aurore Castanié, Jean-François Mercier.

This work has been done during the post-doc of Aurore Castanié, in collaboration with Agnès Maurel from Institut Langevin at ESPCI and Simon Felix from the LAUM (Laboratoire d'Acoustique de l'Université du Maine). Electromagnetic or acoustic metamaterials can be described in terms of equivalent effective, in general anisotropic, media and several techniques exist to determine the effective permeability and permittivity (or effective mass density and bulk modulus in the context of acoustics). Among these techniques, retrieval methods use the measured scattering coefficients for waves incident on a metamaterial slab containing few unit cells. Until now, anisotropic effective slabs have been considered in the literature but they are limited to the case where one of the axes of anisotropy is aligned with the slab interface. We propose an extension to arbitrary orientations of the principal axes of anisotropy and oblique incidence. The retrieval method is illustrated in the electromagnetic case for layered media, and in the acoustic case for array of tilted elliptical particles.

6.2. Wave propagation in heterogeneous media

6.2.1. High order transmission conditions between homogeneous and homogenized periodic half-spaces

Participants: Sonia Fliss, Valentin Vinoles.

This work is a part of the PhD of Valentin Vinoles, and is done in collaboration with Xavier Claeys (LJLL, Paris VI). It is motivated by the fact that classical homogenization theory poorly takes into account interfaces, which is particularly unfortunate when considering negative materials, because important phenomena arise precisely at their surface (plasmonic waves for instance). To overcome this limitation, we want to construct high order transmission conditions. Using matched asymptotics, we have treated the case of a plane interface between a homogeneous and a homogenized periodic half space. The analysis is based on an original combination of Floquet-Bloch transform and a periodic version of Kondratiev techniques. The obtained conditions involve Laplace- Beltrami operators at the interface and requires to solve cell problems in infinite strips.

6.2.2. Multiple scattering by small homogeneities

Participants: Patrick Joly, Simon Marmorat.

This is the topic of the PhD of Simon Marmorat, done in collaboration with the CEA-LIST and with Xavier Claeys (LJLL, Paris VI). We aim at developing an efficient numerical approach to simulate the propagation of waves in concrete, which is modelled as a smooth background medium, with many small embedded heterogeneities. This kind of problem is very costly to handle with classical numerical methods, due the refined meshes needed around the inclusions. To overcome these issues, two models have been developed, which rely on the asymptotic analysis of the problem: each of them can be interpreted as a full space wave equation, which can be discretized using a defects-free mesh, coupled to some auxiliary unknowns accounting for the presence of the inclusions. While the first model is established by using a special Galerkin approximation in the vicinity of the inclusions, the second model only focuses on the far field. The challenge is then to simulate source points coupled to the incident field and this is achieved thanks to the introduction of a special relaxed version of the Dirac mass. Rigorous error estimates as well as some numerical tests have been established, highlighting the efficiency of the two methods.

6.2.3. Finite Element Heterogeneous Multiscale Method for Maxwell's Equations

Participants: Patrick Ciarlet, Sonia Fliss, Christian Stohrer.

This work is the subject of the post-doc of Christian Stohrer. The standard Finite Element Heterogeneous Multiscale Method (FE-HMM) can be used to approximate the effective behavior of solutions to the classical Helmholtz equation in highly oscillatory media. Using a novel combination of well-known results about FE-HMM and the notion of T-coercivity, we derive an a priori error bound. Numerical experiments corroborate the analytical findings. We work now on the application of HMM in presence of interfaces, for Maxwell's equations and finally in presence of high contrast materials.

6.2.4. Effective boundary conditions for strongly heterogeneous thin layers

Participants: Matthieu Chamaillard, Patrick Joly.

This topic is the object of the PhD of Matthieu Chamaillard, done in collaboration with Houssem Haddar (CMAP École Polytechnique). We are interested in the construction of effective boundary conditions for the diffraction of waves by an obstacle covered with a thin coating whose physical characteristics vary "periodically". The width of the coating and the period are both proportional to the same small parameter δ. In the scalar case, we proved that the error between the exact model (with the thin coat) and the one with the effective boundary condition of order n for $n \in \{1, 2\}$ is of the order $\mathcal{O}(\delta^{n+1})$. This has been checked numerically for some two dimensional configurations. Recently, we also succeeded to extend our theoretical work to Maxwell equations. We found a first order boundary condition of the form $E \times n = \delta ik\mathcal{Z}_\Gamma\left(n \times (H \times n)\right)$ where n is the unit outward normal to the boundary Γ and $\mathcal{Z}_\Gamma$ is a second order tangential differential operator along Γ. The coefficients of this operator depend only on the deformation mapping ψ_Γ and the material properties of the coating, through the resolution of particular unbounded cell problems in the flat reference configuration. When the coating is homogeneous, one recovers the well known first order thin layer condition. We have moreover proven that this effective condition provides an error of the order $\mathcal{O}(\delta^2)$.

6.3. Spectral theory and modal approaches for waveguides

6.3.1. Guided modes in ladder-like open periodic waveguides

Participants: Sonia Fliss, Patrick Joly, Elizaveta Vasilevskaya.

This work is done in the context of the PhD of Elizaveta Vasilevskaya, in collaboration with Bérangère Delourme, from Paris 13 University. We consider the theoretical and numerical aspects of the wave propagation in ladder-like periodic structures. We exhibit situations where the introduction of a lineic defect into the geometry of the domain leads to the appearance of guided modes and we provide numerical simulations to illustrate the results. From the theoretical point of view, the problem is studied by asymptotic analysis methods, the small parameter being the thickness of the domain, so that when the thickness of the structure is small enough, the domain approaches a graph. Numerical computations are based on specific transparent conditions for periodic media.

6.3.2. Absence of trapped modes for a class of unbounded propagative media

Participants: Anne-Sophie Bonnet-Ben Dhia, Christophe Hazard, Sonia Fliss, Antoine Tonnoir.

We have proposed a new approach to prove that there does not exist square-integrable solutions to the two-dimensional Helmholtz equation in a homogeneous conical domain with a vertex angle greater than π. This shows that for a medium filling the whole plane, there can be no trapped modes if all the inhomogeneities (penetrable or not) are concentrated in a conical domain with a vertex angle less than π. The proof uses the compatibility of Fourier representations of the field in different half-spaces. One interesting consequence of our result concerns the case of curved open waveguides (e.g., bended optical fibers). Unlike closed waveguides for which trapped modes confined near the bend may occur, our result implies that trapped modes cannot exist if the core of the waveguide is located in a cone with vertex angle less than π. Our results can be extended to higher space dimensions, and to some Y-junctions of open waveguides (using a generalized Fourier transform instead of the usual one).

6.3.3. Reduced graph models for networks of thin co-axial electromagnetic cables

Participants: Geoffrey Beck, Patrick Joly.

This work is the object of the PhD of Geoffrey Beck and is done collaboration with Sébastien Imperiale (Inria, MEDISIM). The general context is the non destructive testing by reflectometry of electric networks of co-axial cables with heterogeneous cross section and lossy materials, which is the subject of the ANR project SODDA. We consider electromagnetic wave propagation in a network of thin coaxial cables (made of a dielectric material which surrounds a metallic inner-wire). The goal is to reduce $3D$ Maxwell's equations to a quantum graph in which, along each edge, one is reduced to compute the electrical potential and current by solving 1D wave equations (the telegrapher's model) coupled by vertex conditions. Using the method of matched asymptotics, we have derived and justified improved Kirchhoff conditions.

6.3.4. Geometrical transformations for waveguides of complex shapes

Participant: Jean-François Mercier.

In collaboration with Agnès Maurel from the Langevin Institut and Simon Felix from the LAUM, we have developed multimodal methods to describe the acoustic propagation in rigid waveguides of general shapes, with varying curvature and cross section. A key feature is the use of a flexible geometrical transformation to a virtual space in which the waveguide is straight but associated to Robin boundary conditions. We have revisited an efficient method developed earlier which consists in adding two extra non-physical modes to the usual modal expansion of the field on the Neumann guided modes, in order to obtain a better convergence of the modal series.

This method has been extended to a half guide with an end wall of general shape, transformed into a flat surface by a geometrical transformation, thus avoiding to question the Rayleigh hypothesis. The transformation only affects a bounded inner region that naturally matches the outer region, which allows to easily select the ingoing and outgoing waves.

6.4. Inverse problems

6.4.1. Quasi-Reversibility method and exterior approach for evolution problems

Participants: Eliane Bécache, Laurent Bourgeois.

This work is a collaboration with Jérémi Dardé from Toulouse University and has been the object of the internship of Lucas Franceschini, student at ENSTA. We address some linear ill-posed problems involving the heat or the wave equation, in particular the backward heat equation and the heat/wave equation with lateral Cauchy data. The main objective is to introduce some variational mixed formulations of quasi-reversibility which enable us to solve these ill-posed problems by using classical Lagrange finite elements. We have also designed a new approach called the “exterior approach” to solve inverse obstacle problems with initial condition and lateral Cauchy data for heat/wave equation. It is based on a combination of an elementary level set method and the quasi-reversibility methods we have just mentioned. Some numerical experiments have proved the feasibility of our strategy in all those situations.

6.4.2. Uniqueness and non-uniqueness results for the inverse Robin problem

Participant: Laurent Bourgeois.

This work is a collaboration with Laurent Baratchart and Juliette Leblond (Inria, APICS). We consider the classical Robin inverse problem, which consists in finding the ratio between the normal derivative and the trace of the solution (the Robin coefficient) on a subset of the boundary, given the Cauchy data (both the normal derivative and the trace of the solution) on the complementary subset. More specifically, we consider a Robin coefficient which is merely in L^∞ and a Neumann data in L^2. In the $2D$ case we prove uniqueness of the Robin coefficient for a problem governed in a Lipschitz domain by a conductivity equation with a conductivity chosen in $W^{1,r}$, where $r > 2$. We also prove a non-uniqueness result in the $3D$ case. In two dimensions, the proof relies on complex analysis, while in higher dimension, the proof relies on a famous counterexample to unique continuation by Bourgain and Wolff.

6.4.3. Higher-order expansion of misfit functional for defect identification in elastic solids

Participants: Marc Bonnet, Rémi Cornaggia.

This work, done in the context of the PhD of Rémi Cornaggia, concerns the defect identification by time-harmonic elastodynamic measurements. We propose a generalization to higher orders of the concept of topological derivative, by expanding the least-squares functional in powers of the small radius of a trial inclusion. This expansion is facilitated by resorting to an adjoint state. With this approach, a region of interest may be exhaustively probed at reasonable computational cost.

6.4.4. Inverse scattering and invisibility with a finite set of emitted-received waves

Participant: Anne-Sophie Bonnet-Ben Dhia.

In collaboration with Lucas Chesnel from CMAP at Ecole Polytechnique and Sergei Nazarov from Saint-Petersburg University, we investigate a time harmonic acoustic scattering problem by a compactly supported penetrable inclusion in the free space. We consider cases where an observer can produce incident plane waves and measure the far field pattern of the resulting scattered field only in a finite set of directions. In this context, we say that a wavenumber is a non-scattering wavenumber if the associated relative scattering matrix has a non trivial kernel. Under certain assumptions on the physical coefficients of the inclusion, we have shown that the non-scattering wavenumbers form a (possibly empty) discrete set. Then, for a given real wavenumber, we built a constructive technique (which provides a numerical algorithm) to prove that there exist inclusions for which the corresponding relative scattering matrix is null. These inclusions have the important property to be impossible to detect from far field measurements.

6.4.5. Energy-based cost functional for three-dimensional transient elastodynamic imaging

Participant: Marc Bonnet.

This work is a collaboration with Wilkins Aquino (Duke University, USA). It is concerned with large-scale three-dimensional inversion under transient elastodynamic conditions by means of the modified error in constitutive relation (MECR), an energy-based, cost functional. Each evaluation of a time-domain MECR cost functional involves the solution of two elastodynamic problems (one forward, one backward), which moreover are coupled (unlike the case of L^2 misfit functionals). This coupling creates a major computational bottleneck, making MECR-based inversion difficult for spatially 2D or 3D configurations. To overcome this obstacle, we propose an approach whose main ingredients are (a) setting the entire computational procedure in a consistent time-discrete framework that incorporates the chosen time-stepping algorithm, and (b) using an iterative successive over-relaxation-like method for the resulting stationarity equations. The resulting MECR-based inversion algorithm is formulated under quite general conditions, allowing for 3D transient elastodynamics, straightforward use of available parallel solvers, a wide array of time-stepping algorithms commonly used for transient structural dynamics, and flexible boundary conditions and measurement settings. The proposed MECR algorithm is then demonstrated on computational experiments involving 2D and 3D transient elastodynamics and up to over 500 000 unknown elastic moduli.

6.5. Integral equations

6.5.1. Fast solution of the BEM system in 3-D frequency-domain elastodynamics

Participants: Stéphanie Chaillat, Patrick Ciarlet, Luca Desiderio.

The main advantage of the Boundary Element Method (BEM) is that only the domain boundaries are discretized leading to a drastic reduction of the total number of degrees of freedom. In traditional BE implementation the dimensional advantage with respect to domain discretization methods is offset by the fully-populated nature of the BEM coefficient matrix. Using the $\mathcal{H}$-matrix arithmetic and low-rank approximations (performed with Adaptive Cross Approximation), we derive a fast direct solver for the BEM system in 3-D frequency-domain elastodynamics. We assess the numerical efficiency and accuracy on the basis of numerical results obtained for problems having known solutions. In particular, we study the efficiency of low-rank approximations when the frequency is increased. The efficiency of the method is also illustrated to study seismic wave propagation in 3-D domains. This is done in partnership with SHELL company in the framework of the PhD of Luca Desiderio.

6.5.2. OSRC preconditioner for 3D elastodynamics

Participant: Stéphanie Chaillat.

This work is done in collaboration with Marion Darbas from University of Picardie and Frédérique Le Louer from Technological University of Compiègne. The fast multipole accelerated boundary element method (FM-BEM) is a possible approach to deal with scattering problems of time-harmonic elastic waves by a three-dimensional rigid obstacle. In 3D elastodynamics, the FM-BEM has been shown to be efficient with solution times of order $O(N \log N)$ per iteration (where N is the number of BE degrees of freedom). However, the number of iterations in GMRES can significantly hinder the overall efficiency of the FM-BEM. To reduce the number of iterations, we propose a clever integral representation of the scattered field which naturally incorporates a regularizing operator. When considering Dirichlet boundary value problems, the regularizing operator is a high-frequency approximation to the Dirichlet-to-Neumann operator, and is constructed in the framework of the On-Surface Radiation Condition (OSRC) method. This OSRC-like preconditioner is successfully applied to Dirichlet exterior problems in 3D elastodynamics.

6.5.3. Boundary Integral Formulations for Modeling Eddy Current Testing

Participants: Marc Bonnet, Audrey Vigneron.

This work was a part of the PhD thesis of Audrey Vigneron, and has been done in collaboration with Edouard Demaldent from CEA-List. It concerns the simulation of eddy current non-destructive testing, which aims to assess the presence of defects (cut, corrosion ...) in a conductive, and possibly magnetic, medium. We propose a simple block-SOR solution method for the PMCHWT-type Maxwell integral formulation, that is well suited for the low-frequency, high-conductivity limit typical of eddy current testing methods. We also derive an asymptotic expansion of the Maxwell integral formulation in powers of some relevant (small) non-dimensional number and show its relation to Hiptmair's eddy current integral formulation. Both aspects are validated on 3D numerical experiments.

6.6. Domain decomposition methods

6.6.1. Transparent boundary conditions with overlap in elastic waveguides

Participants: Anne-Sophie Bonnet-Ben Dhia, Sonia Fliss, Antoine Tonnoir.

This work is a part of the PhD of Antoine Tonnoir and is done in partnership with Vahan Baronian form CEA-LIST. We have conceived new transparent boundary conditions for the time-harmonic diffraction problem in an acoustic or elastic waveguide. These new conditions use the natural modal decomposition in the waveguide and are said “with overlap” by analogy with the domain decomposition methods. Among their main advantages, they can be implemented in general elastic anisotropic waveguides, for which usual Dirichlet to Neumann maps are not available. Moreover, the traditional benefit of the overlap for iterative resolution is obtained, independently of the size of the overlap.

6.6.2. Electromagnetic scattering by objects with multi-layered dielectric coatings

Participants: Patrick Joly, Matthieu Lecouvez.

This is the object of the PhD thesis of Matthieu Lecouvez in collaboration with the CEA-CESTA and Francis Collino. We are interested in the diffraction of time harmonic electromagnetic waves by perfectly conducting objects covered by multi-layered (possibly thin) dielectric coatings. This problem is computationally hard when the size of the object is large (typically 100 times larger) with respect to the incident wavelength. In such a situation, the idea is to use a domain decomposition method in which each layer would constitute a subdomain. The transmission conditions between the subdomains involve some specific impedance operators in order to achieve a geometric convergence of the method (compared to the slow algebraic convergence obtained with standard Robin conditions). We propose a practical solution that uses approximations of nonlocal integral operators with appropriate Riesz potentials.

6.6.3. Domain Decomposition Methods for the neutron diffusion equation

Participants: Patrick Ciarlet, Léandre Giret.

Studying numerically the steady state of a nuclear core reactor is expensive, in terms of memory storage and computational time. In particular, one must solve the neutron diffusion equation discretized by finite element techniques, totaling millions of unknowns or more, within a loop. Iterating in this loop allows to compute the smallest eigenvalue of the system, which determines the critical, or non-critical, state of the 3D core configuration. This problem fits within the framework of high performance computing so, in order both to optimize the memory storage and to reduce the computational time, one can use a domain decomposition method, which is then implemented on a parallel computer. The definition of an efficent DDM has been recently addressed for conforming meshes. The development of non-conforming, hence more flexible, methods is under way. Since one is dealing with highly heterogeneous configurations, the regularity of the exact solution can be very low, which then deteriorates the convergence rate of the discretized solution to the exact one. Next, the optimization of the eigenvalue loop will be studied.
This topic is developed in partnership with CEA-DEN (Erell Jamelot). Realistic computations are carried out with the APOLLO3 neutronics code.

6.7. Aeroacoustics

6.7.1. Time-harmonic acoustic scattering in a rotationnal flow

Participants: Antoine Bensalah, Patrick Joly, Jean-François Mercier.

This activity is done in the framework of the PhD of Antoine Bensalah, in partnership with EADS. We study the time-harmonic acoustic radiation in a fluid in a general flow which is not curl free, but has restricted vortical areas. The objective is to take into account the complicated coupling between acoustics and hydrodynamics. The Galbrun approach developed previously in 2D is too expensive in terms of degrees of freedom for 3D simulations. As an alternative, we propose to consider instead the Goldstein equations, which are vectorial only in the vortical areas and remain scalar elsewhere. Extending the proof done for the Galbrun equation, it is possible to prove that the Goldstein equations are well-posed in a domain Ω if the flow is Ω-filling (each point of Ω is reached by a streamline coming from the inflow boundary in a finite time). Then we focused on the case of a rotating flow in an annular geometry, which is not Ω-filling and we proved the well-posedness of the problem .

POLSYS Project-Team

6. New Results

6.1. Highlights of the Year

Jointly with Univ. Of Kaiserslautern (C. Eder), we have released a new open source C library for linear algebra dedicated to Gröbner bases computations (see http://www-polsys.lip6.fr/~jcf/Software/index.html). This new library opens the door to high performance applications

- The library is specialized in reducing matrices generated during Gröbner bases computations. Optimizing this reduction step is crucial for the overall computation.
- Our approach takes even more advantage of the very special structure (quasi unit-triangular sparse matrices with patterns in the data)
- We also reduce the number of operations, in a parallel friendly fashion, by changing the order of the operations in the elimination.
- We present experimental results for sequential and parallel computations on NUMA architectures. We also get good scaling up until 32 (non hyper-threaded) cores: we have speed-ups around 14 or 16.

6.2. Fundamental algorithms and structured polynomial systems

6.2.1. Sparse Gröbner Bases

Sparse elimination theory is a framework developped during the last decades to exploit monomial structures in systems of Laurent polynomials. Roughly speaking, this amounts to computing in a *semigroup algebra*, *i.e.* an algebra generated by a subset of Laurent monomials. In order to solve symbolically sparse systems, we introduce *sparse Gröbner bases*, an analog of classical Gröbner bases for semigroup algebras, and we propose sparse variants of the F_5 and FGLM algorithms to compute them.

In the case where the generating subset of monomials corresponds to the points with integer coordinates in a normal lattice polytope $\mathcal{P} \subset \mathbb{R}^n$ and under regularity assumptions, we prove in [19] complexity bounds which depend on the combinatorial properties of $\mathcal{P}$. These bounds yield new estimates on the complexity of solving 0-dim systems where all polynomials share the same Newton polytope (*unmixed case*). For instance, we generalize the bound $\min(n_1, n_2) + 1$ on the maximal degree in a Gröbner basis of a 0-dim. Bilinear system with blocks of variables of sizes (n_1, n_2) to the multihomogeneous case: $n + 2 - \max_i \left(\lceil (n_i + 1)/d_i \rceil\right)$. We also propose a variant of Fröberg's conjecture which allows us to estimate the complexity of solving overdetermined sparse systems.

Moreover, our prototype "proof-of-concept" implementation shows large speed-ups (more than 100 for some examples) compared to optimized (classical) Gröbner bases software.

6.2.2. Gröbner bases for weighted homogeneous systems

Solving polynomial systems arising from applications is frequently made easier by the structure of the systems. Weighted homogeneity (or quasi-homogeneity) is one example of such a structure: given a system of weights $W = (w_1, \cdots, w_n)$, W-homogeneous polynomials are polynomials which are homogeneous w.r.t the weighted degree $\deg_W \left(X_1^{\alpha_1}, \cdots, X_n^{\alpha_n}\right) = \sum w_i \alpha_i$.

Gröbner bases for weighted homogeneous systems can be computed by adapting existing algorithms for homogeneous systems to the weighted homogeneous case. In [29], we show that in this case, the complexity estimate for Algorithm F5 $\left(\binom{n+d_{\max}-1}{d_{\max}}^{\omega}\right)$ can be divided by a factor $\left(\prod w_i\right)^{\omega}$. For zero-dimensional systems, the complexity of Algorithm FGLMnD^{ω} (where D is the number of solutions of the system) can be divided by the same factor $\left(\prod w_i\right)^{\omega}$. Under genericity assumptions, for zero-dimensional weighted homogeneous systems of W-degree $(d_1, \cdots, d_n)$, these complexity estimates are polynomial in the weighted Bézout bound $\prod_{i=1}^n d_i / \prod_{i=1}^n w_i$.

Furthermore, the maximum degree reached in a run of Algorithm F5 is bounded by the weighted Macaulay bound $\sum(d_i - w_i) + w_n$, and this bound is sharp if we can order the weights so that $w_n = 1$. For overdetermined semi-regular systems, estimates from the homogeneous case can be adapted to the weighted case.

We provide some experimental results based on systems arising from a cryptography problem and from polynomial inversion problems. They show that taking advantage of the weighted homogeneous structure yields substantial speed-ups, and allows us to solve systems which were otherwise out of reach.

6.2.3. Computing necessary integrability conditions for planar parametrized homogeneous potentials

Let $V \in \mathbb{Q}(i)(\mathbf{a}_1, \cdots, \mathbf{a}_n)(\mathbf{q}_1, \mathbf{q}_2)$ be a rationally parametrized planar homogeneous potential of homogeneity degree $k \neq -2, 0, 2$. In [12], we design an algorithm that computes polynomial *necessary* conditions on the parameters $(\mathbf{a}_1, \cdots, \mathbf{a}_n)$ such that the dynamical system associated to the potential V is integrable. These conditions originate from those of the Morales-Ramis-Simó integrability criterion near all Darboux points and make use of Gröbner bases algorithms. The implementation of the algorithm allows to treat applications that were out of reach before, for instance concerning the non-integrability of polynomial potentials up to degree 9. Another striking application is the first complete proof of the non-integrability of the *collinear three body problem*.

6.3. Solving Polynomial Systems over the Reals and Applications

6.3.1. Exact algorithms for polynomial optimization

Let $f, f_1, ..., f_s$ be n-variate polynomials with rational coefficients of maximum degree D and let V be the set of common complex solutions of $\mathbf{F} = (f_1, ..., f_s)$. In [7], we give an algorithm which, up to some regularity assumptions on $\mathbf{F}$, computes an *exact* representation of the global infimum $f^{\star}$ of the restriction of the map $x \to f(x)$ to $V \cap \mathbb{R}^n$, i.e. a univariate polynomial vanishing at $f^{\star}$ and an isolating interval for $f^{\star}$. Furthermore, it decides whether $f^{\star}$ is reached and if so, it returns $x^{\star} \in V \cap \mathbb{R}^n$ such that $f(x^{\star}) = f^{\star}$.

This algorithm is *probabilistic*. It makes use of the notion of polar varieties. Its complexity is essentially *cubic* in $(sD)^n$ and linear in the complexity of evaluating the input. This fits within the best known *deterministic* complexity class $D^{O(n)}$.

We report on some practical experiments of a first implementation that is available as a MAPLE package. It appears that it can tackle global optimization problems that were unreachable by previous exact algorithms and can manage instances that are hard to solve with purely numeric techniques. As far as we know, even under the extra genericity assumptions on the input, it is the first probabilistic algorithm that combines practical efficiency with good control of complexity for this problem.

It is known that point searching in basic semialgebraic sets and the search for globally minimal points in polynomial optimization tasks can be carried out using $(s\,d)^{O(n)}$ arithmetic operations, where n and s are the numbers of variables and constraints and d is the maximal degree of the polynomials involved.
Subject to certain conditions, we associate in [2] to each of these problems an intrinsic system degree which becomes in worst case of order $(n\,d)^{O(n)}$ and which measures the intrinsic complexity of the task under consideration.
We design non-uniform deterministic or uniform probabilistic algorithms of intrinsic, quasi-polynomial complexity which solve these problems.

6.3.2. Algorithms for answering connectivity queries

Let $\mathbf{R}$ be a real closed field and $\mathbf{D} \subset \mathbf{R}$ an ordered domain. In [4], we give an algorithm that takes as input a polynomial $Q \in \mathbf{D}[X_1, ..., X_k]$, and computes a description of a roadmap of the set of zeros, $\mathrm{Zer}(Q, \mathbf{R}^k)$, of Q in $\mathbf{R}^k$. The complexity of the algorithm, measured by the number of arithmetic operations in the ordered domain $\mathbf{D}$, is bounded by $D^{O(k\sqrt{k})}$, where $D = \deg(Q) \geq 2$. As a consequence, there exist algorithms for computing the number of semi-algebraically connected components of a real algebraic set, $Z(Q, \mathbf{R}^n)$, whose complexity is also bounded by $D^{O(n\sqrt{n})}$, where $D = \deg(Q) \geq 2$. The best previously known algorithm for constructing a roadmap of a real algebraic subset of $\mathbf{R}^n$ defined by a polynomial of degree D has complexity $D^{O(n^2)}$.

In [36], we provide a probabilistic algorithm which computes roadmaps for smooth and bounded real algebraic sets such that the output size and the running time are polynomial in $(nD)^{n\log(n)}$. More precisely, the running time of the algorithm is essentially subquadratic in the output size. Even under these extra assumptions, it is the first roadmap algorithm with output size and running time polynomial in $(nD)^{n\log(n)}$.

6.3.3. Nearly Optimal Refinement of Real Roots of a Univariate Polynomial

In [33], we consider the following problem. We assume that a real square-free polynomial A has a degree d, a maximum coefficient bitsize τ and a real root lying in an isolating interval and having no nonreal roots nearby (we quantify this assumption). Then we combine the Double Exponential Sieve algorithm (also called the Bisection of the Exponents), the bisection, and Newton iteration to decrease the width of this inclusion interval by a factor of $t = 2^L$. The algorithm has Boolean complexity $O(d^2\tau + dL)$. This substantially decreases the known bound $O(d^3 + d^2L)$. Furthermore we readily extend our algorithm to support the same complexity bound for the refinement of r real roots, for any $r \leq d$, by incorporating the known efficient algorithms for multipoint polynomial evaluation. The main ingredient for the latter ones is an efficient algorithm for (approximate) polynomial division; we present a variation based on structured matrices computation with quasi-optimal Boolean complexity.

6.3.4. Accelerated Approximation of the Complex Roots of a Univariate Polynomial

Highly efficient and even nearly optimal algorithms have been developed for the classical problem of univariate polynomial root-finding, but this is still an area of active research. By combining some powerful techniques developed in this area we devise in [20] new nearly optimal algorithms, whose substantial merit is their simplicity, important for the implementation.

6.3.5. Nearly Optimal Computations with Structured Matrices

In [21], we estimate the Boolean complexity of multiplication of structured matrices by a vector and the solution of nonsingular linear systems of equations with these matrices. We study four basic most popular classes, that is, Toeplitz, Hankel, Cauchy and Vandermonde matrices, for which the cited computational problems are equivalent to the task of polynomial multiplication and division and polynomial and rational multipoint evaluation and interpolation. The Boolean cost estimates for the latter problems have been obtained by Kirrinnis, except for rational interpolation, which we provide now. All known Boolean cost estimates for these problems rely on using Kronecker product. This implies the d-fold precision increase for the d-th degree output, but we avoid such an increase by relying on distinct techniques based on employing FFT. Furthermore we simplify the analysis and make it more transparent by combining the representation of our tasks and algorithms in terms of both structured matrices and polynomials and rational functions. This also enables further extensions of our estimates to cover Trummer's important problem and computations with the popular classes of structured matrices that generalize the four cited basic matrix classes.

6.3.6. Bounds for the Condition Number for Polynomials with Integer Coefficients

In [31], we consider the problem of bounding the condition number of the roots of univariate polynomials and polynomial systems, when the input polynomials have integer coefficients. We also introduce an aggregate version of the condition numbers and we prove bounds of the same order of magnitude as in the case of the condition number of a single root.

6.4. Solving Systems in Finite Fields, Applications in Cryptology and Algebraic Number Theory

6.4.1. Polynomial-Time Algorithms for Quadratic Isomorphism of Polynomials: The Regular Case

Let $\mathbf{f} = (f_1, ..., f_m)$ and $\mathbf{g} = (g_1, ..., g_m)$ be two sets of $m \geq 1$ nonlinear polynomials in $\mathbb{K}[x_1, ..., x_n]$ ($\mathbb{K}$ being a field). In [25], we consider the computational problem of finding – if any – an invertible transformation on the variables mapping $\mathbf{f}$ to $\mathbf{g}$. The corresponding equivalence problem is known as *Isomorphism of Polynomials with one Secret* (`IP1S`) and is a fundamental problem in multivariate cryptography. Amongst its applications, we can cite Graph Isomorphism (`GI`) which reduces to equivalence of cubic polynomials with respect to an invertible linear change of variables, according to Agrawal and Saxena. The main result of our work is a randomized polynomial-time algorithm for solving `IP1S` for quadratic instances, a particular case of importance in cryptography.

To this end, we show that `IP1S` for quadratic polynomials can be reduced to a variant of the classical module isomorphism problem in representation theory. We show that we can essentially *linearize* the problem by reducing quadratic-`IP1S` to test the orthogonal simultaneous similarity of symmetric matrices; this latter problem was shown by Chistov, Ivanyos and Karpinski (ISSAC 1997) to be equivalent to finding an invertible matrix in the linear space $\mathbb{K}^{n\times n}$ of $n \times n$ matrices over $\mathbb{K}$ and to compute the square root in a certain representation in a matrix algebra. While computing square roots of matrices can be done efficiently using numerical methods, it seems difficult to control the bit complexity of such methods. However, we present exact and polynomial-time algorithms for computing a representation of the square root of a matrix in $\mathbb{K}^{n\times n}$, for various fields (including finite fields), as a product of two matrices. Each coefficient of these matrices lie in an extension field of $\mathbb{K}$ of polynomial degree. We then consider `#IP1S`, the counting version of `IP1S` for quadratic instances. In particular, we provide a (complete) characterization of the automorphism group of homogeneous quadratic polynomials. Finally, we also consider the more general *Isomorphism of Polynomials* (`IP`) problem where we allow an invertible linear transformation on the variables *and* on the set of polynomials. A randomized polynomial-time algorithm for solving `IP` when $\mathbf{f} = (x_1^d, ..., x_n^d)$ is presented. From an algorithmic point of view, the problem boils down to factoring the determinant of a linear matrix (*i.e.* a matrix whose components are linear polynomials). This extends to `IP` a result of Kayal obtained for `PolyProj`.

6.4.2. A Polynomial-Time Key-Recovery Attack on MQQ Cryptosystems

In [15], we investigate the security of the family of MQQ public key cryptosystems using multivariate quadratic quasigroups (MQQ). These cryptosystems show especially good performance properties. In particular, the MQQ-SIG signature scheme is the fastest scheme in the ECRYPT benchmarking of cryptographic systems (eBACS). We show that both the signature scheme MQQ-SIG and the encryption scheme MQQ-ENC, although using different types of MQQs, share a common algebraic structure that introduces a weakness in both schemes. We use this weakness to mount a successful polynomial time key-recovery attack. Our key-recovery attack finds an equivalent key using the idea of so-called good keys that reveals the structure gradually. In the process we need to solve a MinRank problem that, because of the structure, can be solved in polynomial-time assuming some mild algebraic assumptions. We highlight that our theoretical results work in characteristic 2 which is known to be the most difficult case to address in theory for MinRank attacks. Also, we emphasize that our attack works without any restriction on the number of polynomials removed from the public-key, that is, using the minus modifier. This was not the case for previous MinRank like-attacks against MQ schemes. From a practical point of view, we are able to break an MQQ-SIG instance of 80 bits security in less than 2 days, and one of the more conservative MQQ-ENC instances of 128 bits security in little bit over 9 days. Altogether, our attack shows that it is very hard to design a secure public key scheme based on an easily invertible MQQ structure.

6.4.3. Algebraic Cryptanalysis of a Quantum Money Scheme – The Noise-Free Case

In [13], we investigate the Hidden Subspace Problem (HSP$_q$) over $\mathbb{F}_q$:

Input : $p_1, ..., p_m, q_1, ..., q_m \in \mathbb{F}_q[x_1, ..., x_n]$ of degree $d \geq 3$ (and $n \leq m \leq 2n$).
Find : a subspace $A \subset \mathbb{F}_q^n$ of dimension $n/2$ (n is even) such that

$$p_i(A) = 0 \ \forall i \in \{1, ..., m\} \text{ and } q_j(A^\perp) = 0 \ \forall j \in \{1, ..., m\},$$

where $A^\perp$ denotes the orthogonal complement of A with respect to the usual scalar product in $\mathbb{F}_q$.

This problem underlies the security of the first public-key quantum money scheme that is proved to be cryptographically secure under a non quantum but classic hardness assumption. This scheme was proposed by S. Aaronson and P. Christiano at STOC'12. In particular, it depends upon the hardness of HSP_2. More generally, Aaronson and Christiano left as an open problem to study the security of the scheme for a general field $\mathbb{F}_q$. We present a randomized polynomial-time algorithm that solves the HSP_q for $q > 2$ with success probability $\approx 1 - 1/q$. So, the quantum money scheme extended to $\mathbb{F}_q$ is not secure. Finally, based on experimental results and a structural property of the polynomials that we prove, we conjecture that there is also a randomized polynomial-time algorithm solving the HSP_2 with high probability. To support our theoretical results, we also present several experimental results confirming that our algorithms are very efficient in practice. We emphasize that Aaronson and Christiano propose a non-noisy and a noisy version of the public-key quantum money scheme. The noisy version of the quantum money scheme remains secure.

6.4.4. Algebraic Algorithms for LWE Problems

In [23], we analyse the complexity of algebraic algorithms for solving systems of linear equations with *noise*. Such systems arise naturally in the theory of error-correcting codes as well as in computational learning theory. More recently, linear systems with noise have found application in cryptography. The *Learning with Errors* (LWE) problem has proven to be a rich and versatile source of innovative cryptosystems, such as fully homomorphic encryption schemes. Despite the popularity of the LWE problem, the complexity of algorithms for solving it is not very well understood, particularly when variants of the original problem are considered. Here, we focus on and generalise a particular method for solving these systems, due to Arora & Ge, which reduces the problem to non-linear but noise-free system solving. Firstly, we provide a refined complexity analysis for the original Arora-Ge algorithm for LWE. Secondly, we study the complexity of applying algorithms for computing Gröbner basis, a fundamental tool in computational commutative algebra, to solving Arora-Ge-style systems of non-linear equations. We show positive and negative results. On the one hand, we show that the use of Gröbner bases yields an exponential speed-up over the basic Arora-Ge approach. On the other hand, we give a negative answer to the natural question whether the use of such techniques can yield a subexponential algorithm for the LWE problem. Under a mild algebraic assumption, we show that it is highly unlikely that such an improvement exists. We also consider a variant of LWE known as BinaryError-LWE introduced by Micciancio and Peikert recently. By combining Gröbner basis algorithms with the Arora-Ge modelling, we show under a natural algebraic assumption that BinaryError-LWE can be solved in subexponential time as soon as the number of samples is quasi-linear. We also derive precise complexity bounds for BinaryError-LWE with $m = O(n)$, showing that this new approach yields better results than best currently-known generic (exact) CVP solver as soon as $m/n \geq 6.6$. More generally, our results provide a good picture of the hardness degradation of BinaryError-LWE for various number of samples.. This addresses an open question from Micciancio and Peikert. Whilst our results do not contradict the hardness results obtained by Micciancio and Peikert, they should rule out BinaryError-LWE for many cryptographic applications. The results in this work depend crucially on the assumption the algebraic systems considered systems are not easier and not harder to solve than a random system of equations. We have verified experimentally such hypothesis. We also have been able to prove formally the assumptions is several restricted situations. We emphasize that these issues are highly non-trivial since proving our assumptions in full generality would allow to prove a famous conjecture in commutative algebra known as Fröberg's Conjecture.

6.4.5. Practical Cryptanalysis of a Public-Key Encryption Scheme Based on New Multivariate Quadratic Assumptions

In [10], we investigate the security of a public-key encryption scheme introduced by Huang, Liu and Yang (HLY) at PKC'12. This new scheme can be provably reduced to the hardness of solving a set of quadratic equations whose coefficients of highest degree are chosen according to a discrete Gaussian distributions. The other terms being chosen uniformly at random. Such a problem is a variant of the classical problem of solving a system of non-linear equations (PoSSo), which is known to be hard for random systems. The main hypothesis of Huang, Liu and Yang is that their variant is not easier than solving PoSSo for random instances. In this paper, we disprove this hypothesis. To this end, we exploit the fact that the new problem proposed by Huang, Liu and Yang reduces to an easy instance of the Learning With Errors (LWE) problem. The main contribution of this paper is to show that security and efficiency are essentially incompatible for the HLY proposal. That is, one cannot find parameters which yield a secure and a practical scheme. For instance, we estimate that a public-key of at least 1.03 GB is required to achieve 80-bit security against the simplest of our attacks. As a proof of concept, we present 3 practical attacks against all the parameters proposed by Huang, Liu and Yang. With the most efficient attack, we have been able to recover the private-key in roughly 5 minutes for the first challenge (i.e. Case 1) proposed by HLY and less than 30 minutes for the second challenge (i.e. Case 2).

6.4.6. Lazy Modulus Switching for the BKW Algorithm on LWE

Some recent constructions based on LWE do not sample the secret uniformly at random but rather from some distribution which produces small entries. The most prominent of these is the binary-LWE problem where the secret vector is sampled from $\{0,1\}^*$ or $\{-1,0,1\}^*$. In [9], we present a variant of the BKW algorithm for binary-LWE and other small secret variants and show that this variant reduces the complexity for solving binary-LWE. We also give estimates for the cost of solving binary-LWE instances in this setting and demonstrate the advantage of this BKW variant over standard BKW and lattice reduction techniques applied to the SIS problem. Our variant can be seen as a combination of the BKW algorithm with a lazy variant of modulus switching which might be of independent interest.

In [1], we present a study of the complexity of the Blum-Kalai-Wasserman (BKW) algorithm when applied to the Learning with Errors (LWE) problem, by providing refined estimates for the data and computational effort requirements for solving concrete instances of the LWE problem. We apply this refined analysis to suggested parameters for various LWE-based cryptographic schemes from the literature and compare with alternative approaches based on lattice reduction. As a result, we provide new upper bounds for the concrete hardness of these LWE-based schemes. Rather surprisingly, it appears that BKW algorithm outperforms known estimates for lattice reduction algorithms starting in dimension $n \approx 250$ when LWE is reduced to SIS. However, this assumes access to an unbounded number of LWE samples.

6.4.7. Algebraic Attack against Variants of McEliece with Goppa Polynomial of a Special Form

In [17], we present a new algebraic attack against some special cases of Wild McEliece Incognito, a generalization of the original McEliece cryptosystem. This attack does not threaten the original McEliece cryptosystem. We prove that recovering the secret key for such schemes is equivalent to solving a system of polynomial equations whose solutions have the structure of a usual vector space. Consequently, to recover a basis of this vector space, we can greatly reduce the number of variables in the corresponding algebraic system. From these solutions, we can then deduce the basis of a GRS code. Finally, the last step of the cryptanalysis of those schemes corresponds to attacking a McEliece scheme instantiated with particular GRS codes (with a polynomial relation between the support and the multipliers) which can be done in polynomial-time thanks to a variant of the Sidelnikov-Shestakov attack. For Wild McEliece & Incognito, we also show that solving the corresponding algebraic system is notably easier in the case of a non-prime base field $\mathbb{F}_q$. To support our theoretical results, we have been able to practically break several parameters defined over a non-prime base field $q \in \{9,16,25,27,32\}, t < 7$, extension degrees $m \in \{2,3\}$, security level up to 2^{129} against information set decoding in few minutes or hours.

6.4.8. Folding Alternant and Goppa Codes with Non-Trivial Automorphism Groups

The main practical limitation of the McEliece public-key encryption scheme is probably the size of its key. A famous trend to overcome this issue is to focus on subclasses of alternant/Goppa codes with a non trivial automorphism group. Such codes display then *symmetries* allowing compact parity-check or generator matrices. For instance, a key-reduction is obtained by taking *quasi-cyclic* (QC) or *quasi-dyadic* (QD) alternant/Goppa codes. We show in [6], [18], [28] that the use of such *symmetric* alternant/Goppa codes in cryptography introduces a fundamental weakness. It is indeed possible to reduce the key-recovery on the original symmetric public-code to the key-recovery on a (much) smaller code that has not anymore symmetries. This result is obtained thanks to a new operation on codes called *folding* that exploits the knowledge of the automorphism group. This operation consists in adding the coordinates of codewords which belong to the same orbit under the action of the automorphism group. The advantage is twofold: the reduction factor can be as large as the size of the orbits, and it preserves a fundamental property: folding the dual of an alternant (*resp*. Goppa) code provides the dual of an alternant (*resp*. Goppa) code. A key point is to show that all the existing constructions of alternant/Goppa codes with symmetries follow a common principal of taking codes whose support is globally invariant under the action of affine transformations (by building upon prior works of T. Berger and A. Dür). This enables not only to present a unified view but also to generalize the construction of QC, QD and even *quasi-monoidic* (QM) Goppa codes. All in all, our results can be harnessed to boost up any key-recovery attack on McEliece systems based on symmetric alternant or Goppa codes, and in particular algebraic attacks.

6.4.9. Rounding and Chaining LLL: Finding Faster Small Roots of Univariate Polynomial Congruences

In a seminal work at EUROCRYPT '96, Coppersmith showed how to find all small roots of a univariate polynomial congruence in polynomial time: this has found many applications in public-key cryptanalysis and in a few security proofs. However, the running time of the algorithm is a high-degree polynomial, which limits experiments: the bottleneck is an LLL reduction of a high-dimensional matrix with extra-large coefficients. We present in [11] the first significant speedups over Coppersmith's algorithm. The first speedup is based on a special property of the matrices used by Coppersmith's algorithm, which allows us to provably speed up the LLL reduction by rounding, and which can also be used to improve the complexity analysis of Coppersmith's original algorithm. The exact speedup depends on the LLL algorithm used: for instance, the speedup is asymptotically quadratic in the bit-size of the small-root bound if one uses the Nguyen-Stehlé L2 algorithm. The second speedup is heuristic and applies whenever one wants to enlarge the root size of Coppersmith's algorithm by exhaustive search. Instead of performing several LLL reductions independently, we exploit hidden relationships between these matrices so that the LLL reductions can be somewhat chained to decrease the global running time. When both speedups are combined, the new algorithm is in practice hundreds of times faster for typical parameters.

6.4.10. Symmetrized summation polynomials: using small order torsion points to speed up elliptic curve index calculus

Decomposition-based index calculus methods are currently efficient only for elliptic curves E defined over non-prime finite fields of very small extension degree n. This corresponds to the fact that the Semaev summation polynomials, which encode the relation search (or "sieving"), grows over-exponentially with n. Actually, even their computation is a first stumbling block and the largest Semaev polynomial ever computed is the 6-th. Following ideas from Faugère, Gaudry, Huot and Renault, our goal is to use the existence of small order torsion points on E to define new summation polynomials whose symmetrized expressions are much more compact and easier to compute. This setting allows to consider smaller factor bases, and the high sparsity of the new summation polynomials provides a very efficient decomposition step. In [16], the focus is on 2-torsion points, as it is the most important case in practice. We obtain records of two kinds: we successfully compute up to the 8-th symmetrized summation polynomial and give new timings for the computation of relations with degree 5 extension fields.

6.4.11. Sub-cubic Change of Ordering for Gröner Basis: A Probabilistic Approach

The usual algorithm to solve polynomial systems using Gröbner bases consists of two steps: first computing the DRL Gröbner basis using the F5 algorithm then computing the LEX Gröbner basis using a change of ordering algorithm. When the Bézout bound is reached, the bottleneck of the total solving process is the change of ordering step. For 20 years, thanks to the FGLM algorithm the complexity of change of ordering is known to be cubic in the number of solutions of the system to solve. We show in [14] that, in the generic case or up to a generic linear change of variables, the multiplicative structure of the quotient ring can be computed with no arithmetic operation. Moreover, given this multiplicative structure we propose a change of ordering algorithm for Shape Position ideals whose complexity is polynomial in the number of solutions with exponent ω where $2 \leq \omega < 2.3727$ is the exponent in the complexity of multiplying two dense matrices. As a consequence, we propose a new Las Vegas algorithm for solving polynomial systems with a finite number of solutions by using Gröbner basis for which the change of ordering step has a sub-cubic (i.e. with exponent ω) complexity and whose total complexity is dominated by the complexity of the F5 algorithm. In practice we obtain significant speedups for various polynomial systems by a factor up to 1500 for specific cases and we are now able to tackle some instances that were intractable.

POMDAPI Project-Team

5. New Results

5.1. A posteriori error estimates

Participant: Martin Vohralík.

In [2], we have been able to derive an a posteriori error estimate for the numerical approximation of the two-phase flow problem. This is a cornerstone model problem for porous media, describing the flow of two immiscible and incompressible fluids. We take into account the capillary pressure, whence the model features such difficulties as coupling of partial differential equations with algebraic constraints, strong nonlinearities, degeneracy (disappearance of the diffusion term), advection dominance and consequent forming of sharp evolving fronts, or highly nonlinear and very badly conditioned systems of algebraic equations. Our analysis covers a large class of spatial discretizations in a unified setting, with fully implicit time stepping. We also show how the different error components, namely the spatial discretization error, the temporal discretization error, the linearization error, and the algebraic solver error can be distinguished and estimated separately. This gives rise to efficient adaptive stopping criteria, enabling to spare many useless iterations. The practical impact of our results is that even for this complicated model problem, the overall error committed in a numerical approximation can be fully controlled and, moreover, the simulation time can be reduced by factors typically of an order of magnitude. This result has then been extended in [4] to the compositional model of multiphase Darcy flow, where an arbitrary number of phases can be present, and where each phase can be composed of several components. Later, in [12], still a possible dependence on the temperature has been added. The last two references also contain convincing numerical illustrations on real-life reservoir engineering examples.

5.2. Optimization

Participants: Jean Charles Gilbert, Émilie Joannopoulos, Cédric Josz.

5.2.1. Polynomial optimization

A polynomial optimization problem (POP) consists in minimizing a multivariate real polynomial on a set K defined by polynomial inequalities and equalities. In its full generality it is a non-convex, multi-extremal, difficult global optimization problem. More than a decade ago, J. B. Lasserre proposed to solve a POP by a hierarchy of convex semidefinite programming (SDP) relaxations of increasing size and precision. Each problem in the hierarchy has a primal SDP formulation (a relaxation of a moment expression of the POP) and a dual SDP formulation (a sum-of-squares polynomial relaxation of the POP). In [18], we show that there is no duality gap between each primal and dual SDP problem in Lasserre's hierarchy, provided one of the constraints in the description of set K is a ball constraint. Our proof uses elementary results on SDP duality and it does not assume that K has a strictly feasible point.

5.2.2. Convex quadratic optimization

Convex quadratic optimization deals with problems consisting in minimizing a convex quadratic function on a polyhedron. In [3], we analyzed the behavior of the augmented Lagrangian algorithm when it deals with an *infeasible* convex quadratic optimization problem; this situation is important to master in order to be able to solve correctly the QPs that are generated by the SQP (or Newton-like) algorithm to solve a nonlinear optimization problem, QPs whose feasibility is not guaranteed. It is shown that the algorithm finds a point that, on the one hand, satisfies the constraints shifted by the smallest possible shift that makes them feasible and, on the other hand, minimizes the objective on the corresponding shifted constrained set. The speed of convergence to such a point is globally linear, with a rate that is inversely proportional to the augmentation parameter. This suggests a rule for determining the augmentation parameter that aims at controlling the speed of convergence of the shifted constraint norm to zero; this rule has the advantage of generating bounded augmentation parameters even when the problem is infeasible. The approach has also been implemented in the pieces of software OQLA and QPALM during the ADT MINOQS (see section 4.2 and [16], [14], [15]).

POPIX Team

6. New Results

6.1. Highlights of the Year

Marc Lavielle published the book, *Mixed Effects Models for the Population Approach: Models, Tasks, Methods and Tools* (Chapman & Hall/CRC), which presents a rigorous framework for describing, implementing, and using mixed effects models. With these models, readers can perform parameter estimation and modeling across a whole population of individuals at the same time.

6.2. New result 1

We have proposed a nonlinear mixed-effects framework to jointly model longitudinal and repeated time-to-event data. A parametric nonlinear mixed-effects model is used for the longitudinal observations and a parametric mixed-effects hazard model for repeated event times. We have shown the importance for parameter estimation of properly calculating the conditional density of the observations (given the individual parameters) in the presence of interval and/or right censoring. Parameters are estimated by maximizing the exact joint likelihood with the Stochastic Approximation Expectation-Maximization algorithm. This workflow for joint models is now implemented in the Monolix software, and illustrated on several simulated and real data examples.

6.3. New result 2

We have succesfully extended the methodologies previously developed for ordinary differential equations (ODE) to delay differential equations (DDE). A C++ solver for DDE, and based on an explicit Runge-Kutta scheme, has been developed. This solver can now be used with Monolix, a platform for population modeling of longitudinal data, MlxPlore, a tool for the exploration of complex models and Simulx a R and Matlab function for the simulation of longitudinal data. We use.

POSTALE Team

5. New Results

5.1. Highlights of the Year

CovTrack: Agile multi-target multi-threaded realtime tracker We have developed and highly optimized a multi-target tracking system based on covariance tracking algorithm. The complexity of the algorithm – connected to the number of features – can be tuned to fit the processor computation power (with/without SIMD). Moreover the features can be also selected from a large set of features to adapt the algorithm to the scene and the nature of tracking (indoor/outdoor, pedestrian/car,). Some software and algorithmic transforms have been also applied to accelerate the code for scalar/SIMD processors. [20]

The Light Speed Labeling (LSL) algorithm is still the world fastest connected component labeling (CCL) algorithm. We have proposed a new benchmark that performs fair comparisons for such a data-dependent algorithm (that involves Union-Find algorithm optimization combined with memory and control flow optimization). We show that thanks to its run-based approach and its line-relative labeling, LSL is intrinsically more efficient that all State-of-the-Art pixel-based algorithms, whatever the memory management.[23]

5.2. Excalibur: An Autonomic Cloud Architecture for Executing Parallel Applications

Participants: Alessandro Ferreira Leite, Claude Tadonki, Christine Eisenbeis, Tainá Raiol, Maria Emilia Walter, Alba Cristina de Melo.

IaaS providers often allow the users to specify many requirements for their applications. However, users without advanced technical knowledge usually do not provide a good specification of the cloud environment, leading to low performance and/or high monetary cost. In this context, the users face the challenges of how to scale cloud-unaware applications without re-engineering them. Therefore, in this paper, we propose and evaluate a cloud architecture, namely Excalibur, to execute applications in the cloud. In our architecture, the users provide the applications and the architecture sets up the whole environment and adjusts it at runtime accordingly. We executed a genomics workflow in our architecture, which was deployed in Amazon EC2. The experiments show that the proposed architecture dynamically scales this cloud-unaware application up to 10 instances, reducing the execution time by 73compared to the execution in the configuration specified by the user.[25]

5.3. A Fine-grained Approach for Power Consumption Analysis and Prediction

Participants: Alessandro Ferreira Leite, Claude Tadonki, Christine Eisenbeis, Alba Cristina de Melo.

Power consumption has became a critical concern in modern computing systems for various reasons including financial savings and environmental protection. With battery powered de-vices, we need to care about the available amount of energy since it is limited. For the case of supercomputers, as they imply a large aggregation of heavy CPU activities, we are exposed to a risk of overheating. As the design of current and future hardware is becoming more and more complex, energy prediction or estimation is as elusive as that of time performance. However, having a good prediction of power consumption is still an important request to the computer science community. Indeed, power consumption might become a common performance and cost metric in the near future. A good methodology for energy prediction could have a great impact on power-aware programming, compilation, or runtime monitoring. In this paper, we try to understand from measurements where and how power is consumed at the level of a computing node. We focus on a set of basic programming instructions, more precisely those related to CPU and memory. We propose an analytical prediction model based on the hypothesis that each basic instruction has an average energy cost that can be estimated on a given architecture through a series of micro-benchmarks. The considered energy cost per operation includes both the overhead of the embedding loop and associated (hardware/software) optimizations. Using these precalculated values, we derive a linear extrapolation model to predict the energy of a given algorithm expressed by means of atomic instructions. We then use three selected appli-cations to check the accuracy of our prediction method by comparing our estimations with the corresponding measurements obtained using a multimeter. We show a 9.48% energy prediction on sorting.[27]

5.4. Automated Code Generation for Lattice Quantum Chromodynamics and beyond

Participants: Denis Barthou, Konstantin Petrov, Olivier Brand-Foissac, Olivier Pène, Gilbert Grosdidier, Michael Kruse, Romain Dolbeau, Christine Eisenbeis, Claude Tadonki.

This is ongoing work on a Domain Specific Language which aims to simplify Monte-Carlo simulations and measurements in the domain of Lattice Quantum Chromodynamics. The tool-chain, called Qiral, is used to produce high-performance OpenMP C code from LaTeX sources. We discuss conceptual issues and details of implementation and optimization. The comparison of the performance of the generated code to the well-established simulation software is also made.[17]

5.5. Switchable Scheduling for Runtime Adaptation of Optimization

Participants: Lénaïc Bagnères, Cédric Bastoul.

Parallel applications used to be executed alone until their termination on partitions of supercomputers: a very static environment for very static applications. The recent shift to multicore architectures for desktop and embedded systems as well as the emergence of cloud computing is raising the problem of the impact of the execution context on performance. The number of criteria to take into account for that pur-pose is significant: architecture, system, workload, dynamic parameters, etc. Finding the best optimization for every context at compile time is clearly out of reach. Dynamic optimization is the natural solution, but it is often costly in execution time and may offset the optimization it is en-abling. In this paper, we present a static-dynamic compiler optimization technique that generates loop-based programs with dynamic auto-tuning capabilities with very low overhead. Our strategy introduces switchable scheduling, a family of program transformations that allows to switch between optimized versions while always processing useful computation. We present both the technique to generate self-adaptive programs based on switchable scheduling and experimental evidence of their ability to sustain high-performance in a dynamic environment.[22]

5.6. Efficient distributed randomized algorithms for solving large dense symmetric indefinite linear systems

Participants: Marc Baboulin, Dulceneia Becker, George Bosilca, Anthony Danalis, Jack Dongarra.

Randomized algorithms are gaining ground in high-performance computing applications as they have the potential to outperform deterministic methods, while still providing accurate results. We propose a randomized solver for distributed multicore architectures to efficiently solve large dense symmetric indefinite linear systems that are encountered, for instance, in parameter estimation problems or electromagnetism simulations. Our contribution is to propose efficient kernels for applying random butterfly transformations (RBT) and a new distributed implementation combined with a runtime (PaRSEC) that automatically adjusts data structures, data mappings, and the scheduling as systems scale up. Both the parallel distributed solver and the supporting runtime environment are innovative. To our knowledge, the randomization approach associated with this solver has never been used in public domain software for symmetric indefinite systems. The underlying runtime framework allows seamless data mapping and task scheduling, mapping its capabilities to the underlying hardware features of heterogeneous distributed architectures. The performance of our software is similar to that obtained for symmetric positive definite systems, but requires only half the execution time and half the amount of data storage of a general dense solver. [15]

5.7. Solvers for 3D incompressible Navier-Stokes equations on hybrid CPU/GPU systems

Participants: Yushan Wang, Marc Baboulin, Karl Rupp, Olivier Le Maître, Yann Fraigneau.

We developed a hybrid multicore/GPU solver for the incompressible Navier-Stokes equations with constant coefficients, discretized by the finite difference method. By applying the prediction-projection method, the Navier-Stokes equations are transformed into a combination of Helmholtz-like and Poisson equations for which we describe efficient solvers. We propose a new implementation that takes advantage of GPU accelerators. We present numerical experiments on a current hybrid machine.

5.8. The Numerical Template toolbox: A Modern C++ Design for Scientific Computing

Participants: Pierre Esterie, Joël Falcou, Mathias Gaunard, Jean-Thierry Lapresté, Lionel Lacassagne.

The design and implementation of high level tools for parallel programming is a major challenge as the complexity of modern architectures increases. Domain Specific Languages (or DSL) have been proposed as a solution to facilitate this design but few of those DSL s actually take full advantage of said parallel architectures. In this paper, we propose a library-based solution by designing a C++ DSL s using generative programming: View the MathML source. By adapting generative programming idioms so that architecture specificities become mere parameters of the code generation process, we demonstrate that our library can deliver high performance while featuring a high level API and being easy to extend over new architectures. [18]

5.9. Boost.SIMD: generic programming for portable simdization

Participants: Pierre Esterie, Joël Falcou, Mathias Gaunard, Jean-Thierry Lapresté, Lionel Lacassagne.

Abstract SIMD extensions have been a feature of choice for processor manufacturers for a couple of decades. Designed to exploit data parallelism in applications at the instruction level, these extensions still require a high level of expertise or the use of potentially fragile compiler support or vendor-specific libraries. While a large fraction of their theoretical accelerations can be obtained using such tools, exploiting such hardware becomes tedious as soon as application portability across hardware is required. In this paper, we describe Boost.SIMD, a C++ template library that simplifies the exploitation of SIMD hardware within a standard C++programming model. Boost.SIMD provides a portable way to vectorize computation on Altivec, SSE or AVX while providing a generic way to extend the set of supported functions and hardwares. We introduce a C++standard compliant interface for the users which increases expressiveness by providing a high-level abstraction to handle SIMD operations, an extension-specific optimization pass and a set of SIMD aware standard compliant algorithms which allow to reuse classical C++ abstractions for SIMD computation. We assess Boost.SIMD performance and applicability by providing an implementation of BLAS and image processing algorithms.

5.10. Automatic Task-based Code Generation for High Performance Domain Specific Embedded Language

Participants: Antoine Tran Tan, Joël Falcou, Daniel Etiemble, Harmut Kaiser.

Providing high level tools for parallel programming while sustaining a high level of performance has been a challenge that techniques like Domain Specific Embedded Languages try to solve. In previous works, we investigated the design of such a DSEL-NT2- providing a Matlab-like syntax for parallel numerical computations inside a C++ library. In this paper, we show how NT2 has been redesigned for shared memory systems in an extensible and portable way.[28]

5.11. High Level Transforms for SIMD and low-level computer vision algorithms

Participants: Lionel Lacassagne, Daniel Etiemble, Alain Dominguez, Pascal Vezolle.

This paper presents a review of algorithmic transforms called High Level Transforms for IBM, Intel and ARM SIMD multi-core processors to accelerate the implementation of low level image processing algorithms. We show that these optimizations provide a significant acceleration. A first evaluation of 512-bit SIMD XeonPhi is also presented. We focus on the point that the combination of optimizations leading to the best execution time cannot be predicted, and thus, systematic benchmarking is mandatory. Once the best configuration is found for each architecture, a comparison of these performances is presented. The Harris points detection operator is selected as being *representative* of low level image processing and computer vision algorithms. Being composed of five convolutions, it is more complex than a simple filter and enables more opportunities to combine optimizations. The presented work can scale across a wide range of codes using 2D stencils and convolutions. Such High Level Transforms provide a speedup of x89 on a 2×4 core Intel Xeon processor versus a code that is already SIMDized and OPenMPized.[26]

5.12. What Is the World's Fastest Connected Component Labeling Algorithm?

Participants: Laurent Cabaret, Lionel Lacassagne.

Optimizing connected component labeling is cur- rently a very active research field. Some teams claim to have design the fastest algorithm ever designed. This paper presents a review of these algorithms and a enhanced benchmark that improve classical random images benchmark with a varying granularity set of random images in order to become closer to natural image behavior. Our algorithm, the Light Speed Labeling is from $\times 3.5$ up to $\times 5.3$ faster than the best State-of-the-Art competitor.[23]

5.13. Covariance tracking: architecture optimizations for embedded systems

Participants: Andrés Romero, Lionel Lacassagne, Michèle Gouiffès, Ali Hassan Zahraee.

Covariance matching techniques have recently grown in interest due to their good performances for object retrieval, detection, and tracking. By mixing color and texture information in a compact representation, it can be applied to various kinds of objects (textured or not, rigid or not). Unfortunately, the original version requires heavy computations and is difficult to execute in real time on embedded systems. This article presents a review on different versions of the algorithm and its various applications; our aim is to describe the most crucial challenges and particularities that appeared when implementing and optimizing the covariance matching algorithm on a variety of desktop processors and on low-power processors suitable for embedded systems. An application of texture classification is used to compare different versions of the region descriptor. Then a comprehensive study is made to reach a higher level of performance on multi-core CPU architectures by comparing different ways to structure the information, using single instruction, multiple data (SIMD) instructions and advanced loop transformations. The execution time is reduced significantly on two dual-core CPU architectures for embedded computing: ARM Cortex-A9 and Cortex-A15 and Intel Penryn-M U9300 and Haswell-M 4650U. According to our experiments on covariance tracking, it is possible to reach a speedup greater than 2 on both ARM and Intel architectures, when compared to the original algorithm, leading to real-time execution. [20]

POTIOC Project-Team

6. New Results

6.1. Highlights of the Year

- Acceptance of the ANR project "ISAR" (Interacting with Spatial Augmented Reality) lead by Martin Hachet (Potioc)
- Publication of "Teegi" (Tangible EEG Interface) at UIST14 [15] and more than 13000 views on vimeo until December 2014 (http://vimeo.com/potioc/teegi)

6.2. Teegi -Tangible EEG Interface- and MindMirror for interactive visualization of brain activities

Participants: Jérémy Frey, Renaud Gervais, Fabien Lotte, Martin Hachet.

Typical brain activity visualization tools are usually hard to understand and interpret for novice users. With advances in neurotechnologies (notably BCI) and HCI/AR, we explored the design of new ways to visualize our own brain activity in real-time, for which we proposed two new systems.

We designed Teegi, a Tangible EEG Interface that enables novice users to get to know more about something as complex as brain signals, in an easy, engaging and informative way [15]. To this end, we have designed a new system based on a unique combination of spatial augmented reality, tangible interaction and real-time neurotechnologies (see Figure 5). With Teegi, a user can visualize and analyze his or her own brain activity in real-time, on a tangible character that can be easily manipulated, and with which it is possible to interact. Users can also reveal some specific EEG phenomenons (e.g., sensorimotor rhythms) still using a tangible approach by placing dedicated "mini-teegi" (small pupets) in a designated area on the interaction zone. The whole system has been designed with educational psychology tools in mind to ensure an efficient learning. An explorative study has shown that interacting with Teegi seems to be easy, motivating, reliable and informative. Overall, this suggests that Teegi is a promising and relevant training and mediation tool for the general public.

In addition, together with colleagues from Inria Rennes (team Hybrid), we introduced a novel augmented reality paradigm called ”the Mind-Mirror” which enables such an experience of seeing ”through your own head”, visualizing your brain ”in action and in-situ” [23]. Our approach relies on the use of a semi-transparent mirror positioned in front of a computer screen. A virtual brain is displayed on screen and automatically follows the head movements thanks to an optical face-tracking system. The brain activity is extracted and processed in real-time thanks to an EEG cap wore by the user. A rear view is also proposed thanks to an additional web-cam recording the rear of user’s head (see Figure 6).

6.3. Interaction in mobile augmented reality

Participants: Asier Marzo, Benoît Bossavit, Martin Hachet.

Nowadays, handheld devices are capable of displaying augmented environments in which virtual content overlaps reality. To interact with these environments it is necessary to use a manipulation technique. The objective of a manipulation technique is to define how the input data modify the properties of the virtual objects. Current devices have multi-touch screens that can serve as input. Additionally, the position and rotation of the device can also be used as input creating both an opportunity and a design challenge. In this project we compared three manipulation techniques which namely employ multi-touch, device position and a combination of both. A user evaluation on a docking task revealed that combining multi- touch and device movement yields the best task completion time and efficiency. Nevertheless, using only the device movement and orientation is more intuitive and performs worse only in large rotations. This work has been presented at the ACM Symposium on Spatial User Interaction 2014 [21].

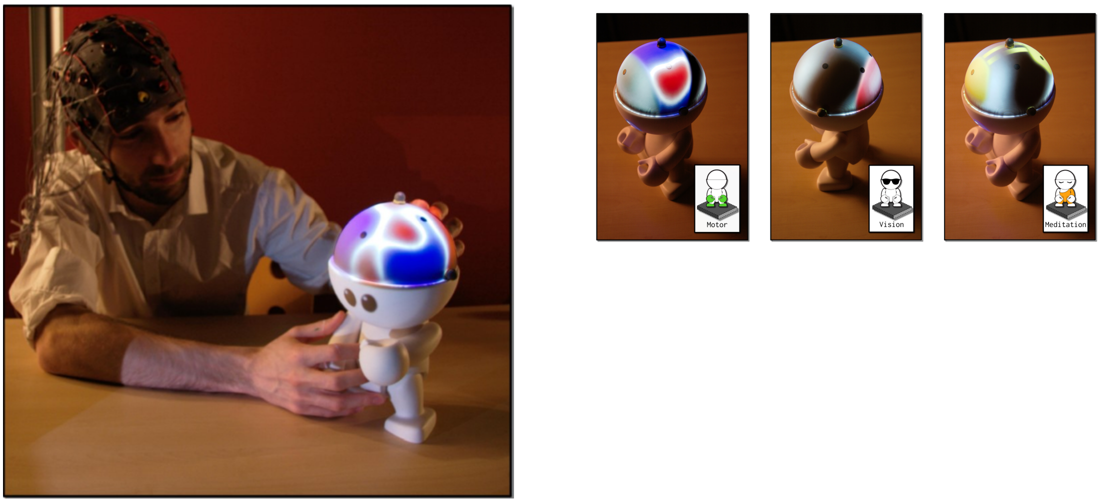


Figure 5. Teegi : a Tangible EEG Interface based on augmented reality.

Figure 6. The Mind Mirror, a new real-time visualization tool of the user's own brain activity based on augmented reality.

Figure 7. Touch (left), tilte (middle) and AR interaction techniques.

In this project we have furthermore evaluated controls based on Augmented Reality (AR), Tilt and Touch for a Point and Shoot Mobile Game (see Figure 7). A user study (n=12) was conducted to compare the three controls in terms of player experience and accuracy. Tilt and AR controls provided more enjoyment, immersion and accuracy to the players than Touch. Nonetheless, touch caused fewer nuisances and was playable under more varied situations. Despite the current technical limitations, we suggest to incorporate AR controls into the mobile games that supported them. Nowadays, AR controls can be implemented on handheld devices as easily as the more established Tilt and Touch controls. However, this study is the first comparison of them and thus its findings could be of interest for game developers. This work has been presented at ISMAR - MASH'D [22].

6.4. CurSAR: Interacting with Spatial Augmented Reality with 2D Input Devices

Participants: Renaud Gervais, Jérémy Frey, Martin Hachet.

Spatial Augmented Reality (SAR) opens interesting perspectives for new generations of mixed reality applications. Compared to traditional HCI contexts, there is little work that studies user performance in SAR. We did an experiment that compared pointing in SAR versus pointing in front of a screen using standard pointing devices (mouse and tablet). The results showed that the users tend to interact in SAR in a way that is similar to the screen condition, without a big loss of performance.

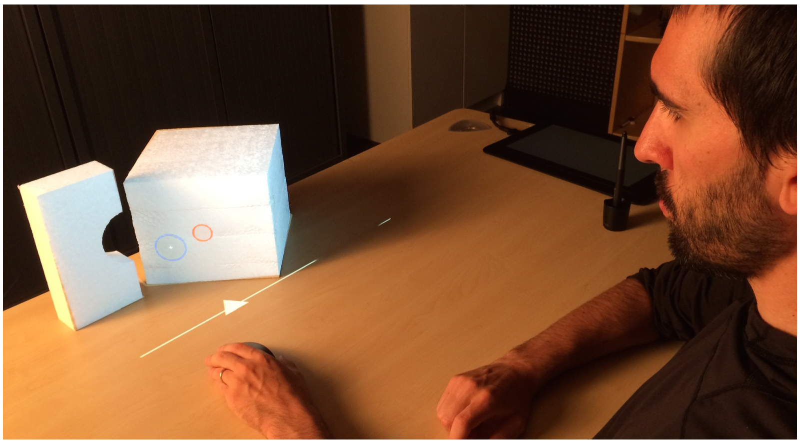


Figure 8. Pointing in spatial augmented reality

6.5. Creative Coding on Objects

Participants: Renaud Gervais, Jérémy Laviole, Asier Marzo, Martin Hachet.

In a near future scenario, we will replace some of our everyday objects with counterparts in form of Computational Objects (COs). COs look similar to the original object; however, inside them there are input sensors, output devices such as displays and a CPU. Furthermore, COs still convey the context and meaning that the original object had. For instance, a clock is associated with time and thus users could expect its CO version to display time-related data. We suggest that any user should be able to easily code new appearances and behaviors for his or her own objects. Using creative coding as a base, we propose to add the notions of affordances and conventions to this programming context. Moreover, we suggest that COs could be used as a creativity support tool although modifying their behavior beyond conventions could confuse the user. Finally, we reckon that with the proper tools, users could also make physical modifications to COs. For example, a retractile cord can be attached to the clock and be used to pull data out and display them in a linear layout.

This work has been presented as a poster at TEI 2014 [30].

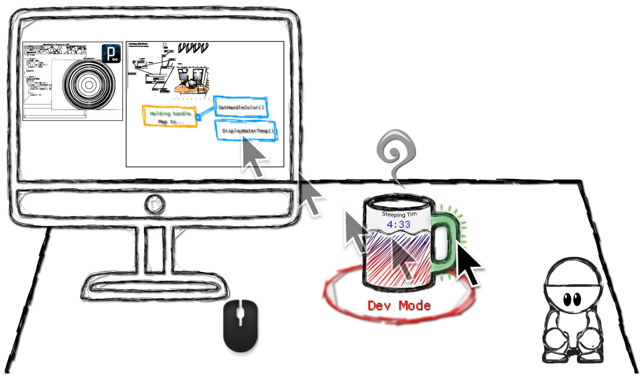


Figure 9. Sketch for a creative coding on objects scenario

6.6. Physiological sensors: bridging human-computer interaction

Participants: Jérémy Frey, Dennis Wobrock, Aurélien Appriou, Christian Mühl, Fabien Lotte, Martin Hachet.

Physiological sensors are not limited to research and medical facilities anymore. They are getting more and more affordable and they become widely accessible to users, as denoted by the popularity of smartphone apps and wearables that track heart rate during fitness activities. Before long, we may see a wide range of sensors embedded into consumer electronic devices. This trend has already started with the arrival of "smartwatches" that could – among other things – detect users' heart beats covertly.

We anticipated this opportunity in order to increase engagement in human-computer interaction, more specifically in human-agent interaction. In [14] we demonstrated that we could increase the social presence of embodied agents – that is, of virtual beings – by simply mirroring the heart rate of users. The "similarity-attraction" effect induces positive emotions toward persons or things that look like us or react as we do. An agent that is associated to a heart beat at the same pace as the user is found more sympathetic. The "similarity-attraction" effect, applied to physiological computing, could help with little effort to improve the acceptance of embodied agents and robots by the general public. (See Figure 10 for setup).

Furthermore, we have taken advantage of physiological sensors in order to evaluate different sorts of human-computer interfaces prior to their release. First, we showed that we could reliably estimate the user's mental workload levels from his/her EEG signals, across different contexts involving different levels of social stress [10]. Then, based on those results, we used a combination of electrocardiography (measure of heart beats), galvanic skin response (measure of sweat on the skin) and electroencephalography (EEG, measure of brain activity) to assess the workload of users during 3D manipulation tasks. The first preliminary results seem to indicate that we might be able to discriminate the parts of the interaction that provokes a high cognitive stress, hence that needs to be improved. This work is in line with the evaluation of visual comfort. We presented earlier this year a pilot study documenting how different virtual depths could cause different levels of discomfort [17], and how this discomfort translates to EEG activity.

Pervasive technologies and physiological computing may be a key component to bridge the gap that too often keeps dividing machines and general public. We believe that it'll help make computers more enjoyable and more usable.

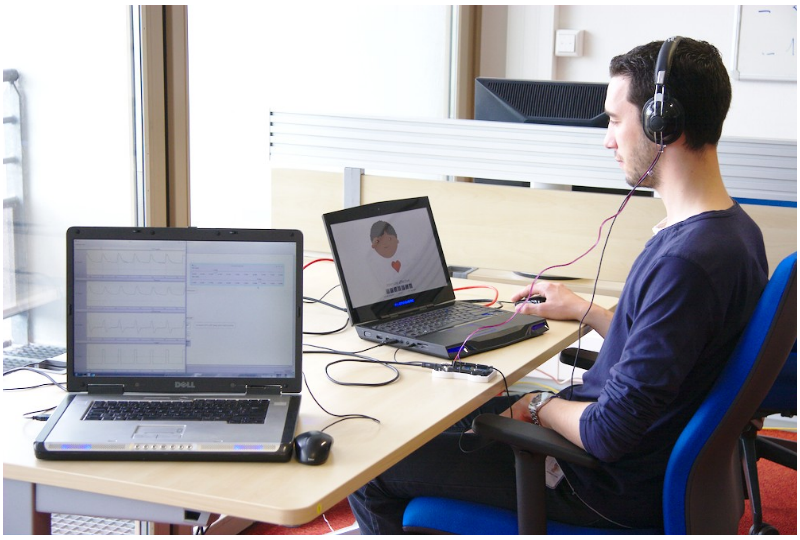


Figure 10. Experimental setup used to study the physiological "similarity-attraction" effect in human-agent interaction.

6.7. Training Approaches for Brain-Computer Interfaces

Participants: Alison Cellard, Martin Hachet, Camille Jeunet, Fabien Lotte, Christian Mühl, Julia Schumacher.

While recent research on Brain-Computer Interfaces (BCI) has highlighted their potential for many applications, they remain barely used outside laboratories due to a lack of robustness. Spontaneous BCI (i.e., mental imagery-based BCI) often rely on mutual learning efforts by the user and the machine, with BCI users learning to produce stable EEG patterns (spontaneous BCI control being widely acknowledged as a skill) while the computer learns to automatically recognize these EEG patterns, using signal processing. Most research so far was focused on signal processing, mostly neglecting the humans in the loop.

Indeed, even if it has been advocated in one of our previous publications (see activity report 2013) that current human training approaches for spontaneous BCI are most likely inappropriate, based on theoretical models, we still needed practical confirmations that users' modest performances at controlling a BCI could be partly due to these inappropriate training protocols. Thus, in our work, we proposed to study standard BCI training protocols without EEG signals, i.e., without a BCI [31]. In particular, we studied how people could learn to do two simple motor tasks using the same training tasks and feedback as the one given to motor imagery BCI users. More precisely, we asked subjects to learn to draw on a graphic tablet a triangle and a circle (the correct size, angles and speed of drawing of these two shapes being unknown to the subject) that can be recognized by the system, using a synchronous training protocol and an extending bar as feedback, like for motor imagery based BCI training. Our results show that most subjects (out of N=20 subjects) improved with this feedback and practice (i.e., the shapes they drew were increasingly more accurately recognized by the system), but that 15% of them completely failed to learn how to draw the correct shapes, despite the simplicity of the motor tasks. This suggests that part of BCI illiteracy/deficiency is likely due to the training protocols currently used.

From the huge variability in users' performances at BCI mastery emerged the following question: Why do some people manage to learn using these protocols and others do not? Our hypothesis here was that these protocols are not adapted to some users' profiles. Thus, we designed an experiment in which we looked for correlations between the personality and cognitive profile of the users and their ability to learn to control a MI-

BCI. Our current results (N=18) show that 1) performances are strongly correlated with users' spatial abilities and 2) we can reliably predict these performances using a model including different psychological factors (like abstractedness, self-reliance or tension). These results are very encouraging as they could lead to reflections about 1) exercices to improve users' spatial abilities and 2) solutions to take into account users' cognitive and personality profiles in BCI training approaches.

Furthermore, it is more and more claimed that visual feedback is not ideal for BCIs as they are conceived for interaction situations in which the visual channel is often overtaxed. Thus, tactile feedback might appear to be more relevant. In order to test this hypothesis, we proposed a study aiming at comparing a standard visual feedback with an equivalent tactile feedback in an appealing training environment containing visual distractors (to mimic an interaction context in which the visual channel is overtaxed). Users had to learn to perform motor-imagery tasks as well as a counting task, and received either a visual or vibrotactile feedback (see Figure 11). Our main result (N=18) is the fact that people receiving tactile feedback perform significantly better (at Motor-Imagery and counting task). This kind of result should encourage the BCI community to replace standard BCI protocols by more motivating training environments and multimodal feedback.

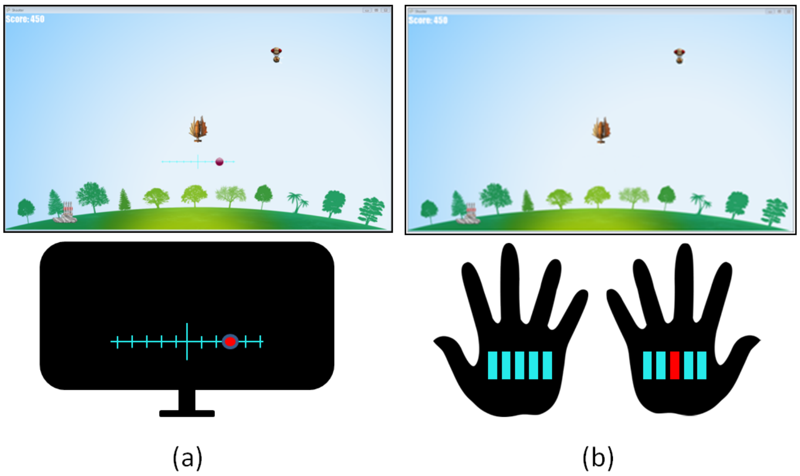


Figure 11. Illustration of the combination of an appealing training environment and a vibrotactile feedback (right, compared to a standard visual feedback, left)

Still regarding the feedback, we explored what kind of information could help the user to perform better mental imagery tasks. As such, we look for physiological features that could predict whether a mental task will be correctly recognized by the BCI, and that could be understood by the user. Among the different features we explored, it appears that the user's relaxation (from a muscular point of view), as measured in EMG activity collected by EEG channels, is one of such features. We are currently building and exploring new BCI training protocols that provide additional information about the user's muscular relaxation as complementary feedback.

6.8. EEG signal processing

Participants: Alison Cellard, Nicoletta Caramia, Fabien Lotte.

Spatial filters are powerful tools for EEG classification for BCI design, able to reduce spatial blurring effects. In particular, optimal spatial filters have been designed to classify EEG signals based on band power features. Unfortunately, there are other relevant EEG features for which no optimal spatial filter exists. This is the case for Phase Locking Value (PLV) features, which measure the synchronization between 2 EEG channels.

Therefore, we proposed to create such a pair of optimal spatial filters for PLV-features [13]. To do so, we optimized a functional measuring the discriminability of PLV-features based on a genetic algorithm. An evaluation of our algorithm on a motor imagery EEG data set showed that using optimized spatial filters led to higher classification performances, and that combining the resulting PLV features with traditional methods boosts the overall BCI performances.

We also wrote a chapter that is an introductory overview and a tutorial of signal processing techniques that can be used to recognize mental states from EEG signals in BCI [26]. More particularly, this chapter presented how to extract relevant and robust spectral, spatial and temporal information from noisy EEG signals (e.g., Band Power features, spatial filters such as Common Spatial Patterns or xDAWN, etc.), as well as a few classification algorithms (e.g., Linear Discriminant Analysis) used to classify this information into a class of mental state. It also briefly touched on alternative, but currently less used approaches.

6.9. Navigation techniques in 3D digital cities on mobile touch devices

Participants: Jacek Jankowski, Thomas Hulin, Martin Hachet.

This project was part of "Villes transparentes" research project in collaboration with Mappy (Solocal group) and Vectuel - VirtuelCity initiated in 2013. It aimed at characterizing today's most common interaction techniques for street-level navigation in 3D digital cities, for mobile touch devices, in terms of their efficiency and usability. To do so, we conducted a user study, where we compared target selection (Go-To), rate control (Joystick), position control, and stroke-based control navigation metaphors (see Figure 12). The results suggest that users performed best with the Go-To interaction technique. The subjective comments showed a preference of novices towards Go-To and expert users towards the Joystick technique. This work has been published at the 3DUI 2014 conference [18].

(a) Target selection (b) Rate control (c) Position control (d) Stroke-based control

Figure 12. Four techniques for navigating in a 3D city on a mobile touch device.)

PRIMA Project-Team

5. New Results

5.1. Highlights of the Year

On March 14, 2014, James Crowley was named Chevalier de l'Ordre national du Mérite.

On August 2014, the paper "Human-Robot Motion: an Attention-Based Navigation Approach" [14] by Thierry Fraichard, Remi Paulin & Patrick Reignier has been nominated for the best paper award at the IEEE Int. Symp. on Robot and Human Interactive Communication (RO-MAN 2014), Edinburgh (UK).

On December 2014, Patrick Reignier was a member of the EDF grand jury for smart energy
BEST PAPER AWARD :
[14] **IEEE Int. Symp. on Robot and Human Interactive Communication (ROMAN)**. T. FRAICHARD, R. PAULIN, P. REIGNIER.

5.2. Attention-Based Navigation

Participants: Thierry Fraichard, Remi Paulin, Patrick Reignier..

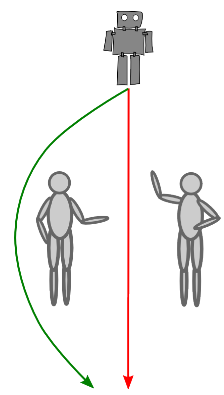


Figure 6. People are not pieces of furniture! Motion in red is definitely shorter but it is not appropriate.

The domain of service-robots is growing fast and has become the focus of many researchers and industrials alike. Their application areas been extremely broad, from logistics to handicap assistance. A large proportion of such robots are expected to share humans' living space and thus must be endowed with navigation capabilities that exceed the standard requirements pertaining to autonomous navigation such as motion safety. In a human populated environment, optimality does not boil down to minimising resources such as time or distance travelled anymore, the robot motion must abide by social/cultural rules and be **appropriate**, *e.g.* Fig. 6 .

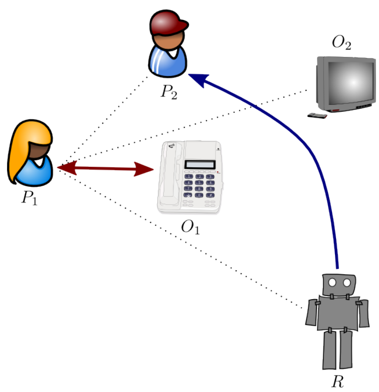


Figure 7. Attention vs activity: although P_1's current activity is being on the phone, part of her attention may be directed towards P_2, the TV set or the robot R. Suppose now that R moves towards P_2 in a way that hides the TV from P_1. Such a behavior would not be appropriate should P_1 be actually paying attention to the TV.

Most of the approaches proposed so far relies upon the definition of so-called **social spaces**, *i.e.* regions in the environment that, for different reasons, the persons consider as psychologically theirs. Such social spaces are primarily characterized using either the position of the person, *e.g.* “Personal space” [42], or the activity it is currently engaged in, *e.g.* “Interaction Space” [47] and “Activity Space” [51]. The most common approach is then to define costmaps on such social spaces: the higher the cost, the less desirable it is for the robot to be at the corresponding position. The costmaps are ultimately used for motion planning and navigation purposes. Such approaches are interesting however their spatial nature (being inside or outside the space) make them less suitable when facing more complicated situations, *e.g.* Fig. 7 . To overcome those limits, we suggest using the psychological concept of **attention**, which plays a central role when humans navigate around each other. Besides lifting the limits of social spaces, this concept brings a new degree of control over the motion of the robot, namely the invasive and distracting character of the robot motion, which have so far proven hard to tackle with the conventional tools such as social spaces. Beside leading appropriate motion, attention-based navigation enable interaction through motion by predicting the quantity of attention the human will give to the robot.

Building upon a computational model of attention that was earlier proposed in [53]. we have developed the novel concept of **attention field**. The attention field is straightforward to define: it is a measure of the amount of attention that a given person would allocate to the robot, should the robot be in a given position/state. It is mapping from the state space of the robot to $\mathbb{R}$. The attention field can serve as an attention predictor that can be used to predict potential attentional situations. This knowledge can in turn be used to decide what the robot should do in the future depending on its current task.

Let us illustrate this on a simple scenario featuring a person, a TV and a robot (denoted by P_1, O_1 and R in Fig. 8 -left). The person is currently watching the TV: this is his current *activity*. This activity relates to his *intention* and is modeled by the yellow vector $\overrightarrow{I}$ in Fig. 8 -left that is directed from the person to the TV. Fig. 8 -right depicts the attention field for the person; it is a mapping from $\mathbb{R}^2$ to $\mathbb{R}$ that gives the amount of attention that the person is paying to the robot when it is at a given position (x, y). Fig. 8 -right should be interpreted as follows: the warmer the color, the higher the amount of attention given by the person to the robot. It integrates both the visual and auditory perception capabilities of P_1.

In 2014, we have furthered the development of the concept of attention field and demonstrated different ways to use its attention prediction capability on various scenarios. The main results obtained have been reported in a conference article that has been nominated for the Best Paper Award [14]. Work is ongoing to quantify

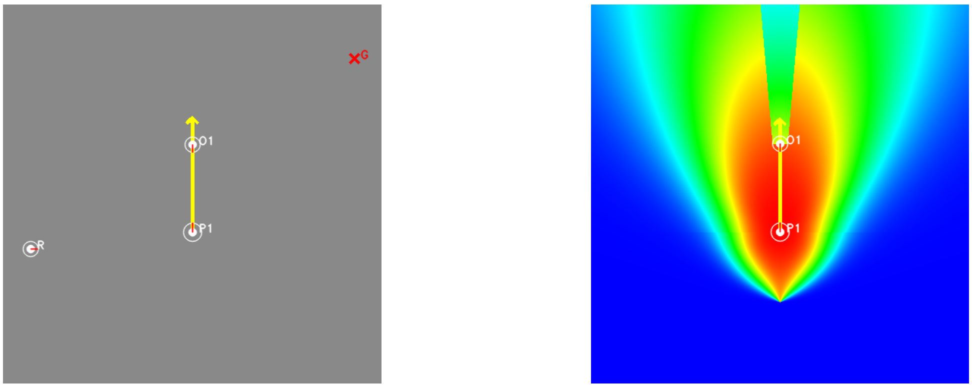


Figure 8. Person-TV-Robot scenario (left), Attention field for the person P_1 (right).

the social "goodness" of the paths provided by our approach, to further the use of the concept of attention on more challenging and dynamic scenarios and to offer an approach to fill the gap between appropriate motion and interaction through motion.

5.3. SPOK: End User Programming for Smart Homes

Participant: Alexandre Demeure.

As part of the CATRENE project AppsGate, we have developed SPOK, an End User Development Environment, that enables inhabitants to control and program their smart Homes via a web interface. The current version of SPOK includes an editor for editing programs using a pseudo-natural language and an interpreter. A multi-syntax editor as well as additional services such as a debugger and a simulator are expected for the second version.

A multi-syntax editor will allow users to build syntactically correct programs using the syntax that is most appropriate to them or by using a combination of them. These syntaxes include pseudo-natural language (i.e. a constrained natural language) and graphical iconic syntax (as exemplified by Scratch [Maloney et al. 2010]). The interaction techniques used to enter programs may be menu-based, free typing, as well as by demonstration in the physical home or by the way of the simulator. The simulator is the dual digital representation of the real home. It is intended to serve also as a debugger for testing and correcting end-user programs.

Whatever syntax used by end-users, programs are translated into syntactic abstract trees whose leaves reference services provided by the Core HMI and/or by the Extended HMI Middleware. The interpreter, executes end-user programs, using the corresponding syntactic abstract trees as input.

In order to support a dynamically extensible grammar as well as to provide end-users with feedforward at the user interface of the editor, the grammar used by the editor is split into 2 parts: the root grammar and the device specific grammars. The root grammar specifies the generic structures of an end-user program: loops, conditions, etc. The device specific grammars are separated from the root grammar to be able to dynamically build the final grammar to be compliant with what is currently installed and detected by the AppsGate server. Each device type brings with it its own events, status and actions. These grammatical elements are injected into the root grammar when generating the parser and for compiling end-user programs.

The language used by end-users to express their programs is a pseudo-natural language using the rule-based programming paradigm. The left hand side of a rule is composed of events and conditions, and the right hand side specifies the actions to be taken when the left hand-side is true or becomes true. A program may include several rules that can be executed either in parallel or sequentially. Once entered, programs are translated into syntactic abstract trees. The interpreter, executes end-user programs, using the corresponding syntactic abstract trees as input. SPOK is implemented as a mix of OSGi and ApAM components where ApAM is in turn a middleware that runs on top of OSGi.

5.4. Qualitative approaches for building energy management

Participant: Patrick Reignier.

Reducing housing energy costs is a major challenge of the 21st century. In the near future, the main issue for building construction is the thermal insulation, but in the longer term, the issues are those of renewable energy (solar, wind, etc.) and smart buildings. Home automation system basically consists of household appliances linked via a communication network allowing interactions for control purposes. Thanks to this network, a load management mechanism can be carried out: it is called distributed control. An optimal home energy management system is still a goal to aim for, because lots of aspects are still not completely fulfilled. Most of the energy systems respect only the energy needs, but they don't tackle the user needs or satisfaction. Energy systems also have a lack when it comes to the dynamicity of the environments (the system ability to adapt).The problem is similar for the existing HMI (Human User Interface) of those Home Automation Systems where only experts can understand the data coming from the sensors and most important, the energy plan coming from management system (How? and Why?). The goal of this study is to propose a house energy model that can be both used to predict at some level energy evolution and that can be understood by the end user. The house energy model is based on Fuzzy Cognitive Maps representing cause-effects relations. It is first designed by an expert and then automatically tuned to a particular house using machine learning approaches. Preliminary experiments have been done this year using the Predis datasets.

PRIVATICS Project-Team

5. New Results

5.1. Highlights of the Year

Vincent Roca was awarded the 3rd Applied Research price of the Fédération des Industries Electriques, Electroniques et Communications (FIEEC), for his transfer activities to the Expway French SME, Lyon, October 8th, 2014.

The team got two major contributions:

- *A Case Study: Privacy Preserving Release of Spatio-temporal Density in Paris* was published by Gergely Acs and Castelluccia at KDD 2014.
- *Censorship in the Wild: Analyzing Internet Filtering in Syria* was published by Chaabane Abdelberi, Mathieu Cunche,and Mohamed Ali Kaafar at IMC 2014.

5.2. Filtering and blocking the Internet

Participants: Mohamed Ali Kaafar, Abdelberi Chaabane, Mathieu Cunche, Cédric Lauradoux, Amrit Kumar.

- **Censorship**

  Based on 600GB leaked logs from appliances used to filter Internet traffic in Syria, we performed an analysis of the Syrian censorship apparatus. This study have been published in ACM Internet Measurement Conference [7].

  We found that the Internet traffic in Syria was filtered in several ways using IP addresses, domain names and keywords. Content sharing, instant messaging and proxy technologies were heavily censored. Some social media such as badoo.com were fully censored, but others such as Facebook are only censored for specific political and religious pages. We also found evidences of successful usage of censorship-circumvention techniques such as Tor and VPN. We also found that P2P file-sharing and Google cache were used to escape censorship blockage.

  While our work might help organizations on both sides of the censorship line, we believe the presented results can help understand the underlying technologies, policies and can inform the design of tools designed to evade the censorship.

- **Attacking filters** Many major Internet companies use probabilistic techniques to filter the users requests or to prevent malicious attacks. In our work [35], [34], we show how they can be polluted/saturated using pre-image attacks and how it increases the false-positive probability. Then, we show how to forge false-positives to mount attacks. In the adversarial settings, we have the liberty to assume that the inputs to the filter are non uniformly distributed. This observation leads to our second contribution: we compute the worst case false-positive probability and obtain new equations for Bloom filter parameters. To support our contributions, we provide four attacks on software applications based on Bloom filter: Bloom-enabled SCRAPY web spider, BITLYDABLOOMS spam filter, SQUID web cache and GOOGLE Safe Browsing. Our attacks retain some form of DoS. They are all based on the forgery of Uniform Resource Locators (URLs) matching certain pre-image or second pre-image property. The impact of our attack ranges from denial-of-service to massively distributed denial-of-service with reflection.

5.3. Selling Off Privacy at Auction

Participants: Claude Castelluccia, Lukasz Olejnik, Minh-Dung Tran.

The first one is a privacy analysis of Real-Time Bidding (RTB) and Cookie Matching (CM). RTB is a technology that allows ad buyers (advertisers) and ad sellers (publishers) to buy and sell ad spaces at real-time auctions through ad exchanges. In RTB, when user visits a publisher page, the ad impression (i.e. one ad display in an ad space) and the user information are immediately broadcast by the ad exchange to a number of bidders (i.e. advertisers or their representatives) for them to bid for the chance to serve ads to this user. CM protocol allows the ad exchange and the bidder to synchronize their cookies of the same user, thus facilitating their exchange of user data.

In [13], we characterize and quantify the potential user web history leakage from ad exchanges to bidders in RTB as a result of exchanging user data. We also discuss and quantify the extent to which companies can potentially collude to increase their tracked user profiles using CM. In addition, we leverage a design characteristic of RTB to observe the winning price of each RTB auction. By analyzing these prices, we show how advertisers evaluate the value of user privacy. This work (titled Selling Off Privacy at Auction) will be presented in NDSS 2014, San Diego, USA in February, 2014.

5.4. Data anonymization

Participants: Claude Castelluccia, Gergely Acs.

With billions of handsets in use worldwide, the quantity of mobility data is gigantic. When aggregated they can help understand complex processes, such as the spread viruses, and built better transportation systems, prevent traffic congestion. While the benefits provided by these datasets are indisputable, they unfortunately pose a considerable threat to location privacy. At KDD 2014 [9], we present a new anonymization scheme to release the spatio-temporal density of Paris, in France, i.e., the number of individuals in 989 different areas of the city released every hour over a whole week. The density is computed from a call-data-record (CDR) dataset, provided by the French Telecom operator Orange, containing the CDR of roughly 2 million users over one week. Our scheme is differential private, and hence, provides provable privacy guarantee to each individual in the dataset. Our main goal with this case study is to show that, even with large dimensional sensitive data, differential privacy can provide practical utility with meaningful privacy guarantee, if the anonymization scheme is carefully designed. This work is part of the national project XData (http://xdata.fr) that aims at combining large (anonymized) datasets provided by different service providers (telecom, electricity, water management, postal service, etc.).

5.5. Wi-Fi and privacy

Participants: Jagdish Achara, Mathieu Cunche, Vincent Roca.

In Android, installing an application implies accepting the permissions it requests, and these permissions are then enforced at runtime. In our WISEC 2014 paper [29], we focus on the privacy implications of the `ACCESS_WIFI_STATE` permission. For this purpose, we analyzed permissions of the 2700 most popular applications on Google Play and found that the `ACCESS_WIFI_STATE` permission is used by 41% of them. We then performed a static analysis of 998 applications requesting this permission and based on the results, chose 88 applications for dynamic analysis. Our analyses reveal that this permission is already used by some companies to collect user Personally Identifiable Information (PII). We also conducted an online survey to study users' perception of the privacy risks associated with this permission. This survey shows that users largely underestimate the privacy implications of this permission. As this permission is very common, most users are therefore potentially at risk.

5.6. Sensor security and privacy

Participant: Marine Minier.

Wireless sensor networks (WSNs) are composed of a large number of low-cost, low-power, and multi-functional sensor nodes that communicate at short distance through wireless links. They are usually deployed in an open and uncontrolled environment where attackers may be present. Due to the use of low-cost materials, hardware components are not tamper-resistant and an adversary could access to a sensor's internal state. With Ochirkhand Erdene-Ochir and Pierre Brunisholz, we continue to work on the notion of resiliency in WSNs [17], [31].

5.7. Buidling blocks

Participant: Marine Minier.

In the context of the BLOC project funded by the ANR, we continue to work on Extended Generalized Feistel Network and on new ligthweight block cipher design (see [30]). We hope to obtain results in this area at the beginning of 2015. With Christine Solnon and Julia Reboul, we work on the formalism of related-key and chosen-key attacks against symmetric key primitives using constraint programming (CP). This preliminary work was presented at the CP 2014 workshop ModRef 2014 in [42].

5.8. Formal and legal issues of privacy

Participants: Thibaud Antignac, Denis Butin, Daniel Le Métayer.

- **Privacy Architectures: Reasoning About Data Minimization and Integrity** Privacy by design will become a legal obligation in the European Community if the Data Protection Regulation eventually gets adopted. However, taking into account privacy requirements in the design of a system is a challenging task. We present an approach based on the specification of privacy architectures at FM 2014 [12] and focus on a key aspect of privacy, data minimisation, and its tension with integrity requirements. We illustrate our formal framework through a smart metering case study.
- **Log Analysis for Data Protection Accountability**

  Accountability is increasingly recognized as a cornerstone of data protection, notably in European regulation, but the term is frequently used in a vague sense. For accountability to bring tangible benefits, the expected properties of personal data handling logs and the assumptions regarding the logging process must be defined with accuracy. At STM 2014 [10], we provide a formal framework for accountability and show the correctness of the log analysis with respect to abstract traces used to specify privacy policies. We also show that compliance with respect to data protection policies can be checked based on logs free of personal data, and describe the integration of our formal framework in a global accountability process.

PROSECCO Project-Team

6. New Results

6.1. Highlights of the Year

This year, we published 17 articles in international peer-reviewed journals and conferences, including papers in prestigious conferences such as POPL (2 papers) and all the top conferences in computer security: IEEE S&P Oakland (2 papers), CRYPTO, ACM CCS, NDSS, and Financial Cryptography. Our papers in these top venues (discussed later in New Results) serve as highlights of our research during the year. In addition to these papers, we published 1 PhD thesis and several technical reports.

We released updates to miTLS, ProVerif, CryptoVerif, and started working on a brand-new version of F*. We discovered serious vulnerabilities in a number of TLS libraries, web browsers, and web servers, resulting in 6 published CVEs, and over a dozen software updates based on our recommendations in widely used software such as Firefox, Chrome, Internet Explorer, Safari, OpenSSL, Java, and Mono.

We organized a winter school "The Joint EasyCrypt-F*-CryptoVerif School 2014" which attracted industrial researchers, academics, and students from around the world. Over 75 students learned to use cryptographic verification tools from instructors at Inria, IMDEA, and Microsoft Research. Two of the tools: CryptoVerif and F* are being developed in collaboration with Inria.

If we were to choose one research theme as our highlight of the year, it would be our activities surrounding Transport Layer Security (TLS):

- At CRYPTO 2014, we published a detailed cryptographic proof of the TLS handshake as implemented in miTLS
- At NDSS 2014, we published a study in the use of X.509 certificates in TLS servers on the web
- At IEEE S&P (Oakland), we published a new attack on the TLS protocol called the *triple handshake*, which affected all TLS libraries and mainstream TLS applications such as web browsers.
- To prevent our attack, we proposed patches to major software libraries as part of responsible disclosure. Our research directly led to security updates for all major web browsers and TLS implementations.
- We also proposed a long-term countermeasure for our attack, the TLS Session Hash extension, which we published as an internet draft and presented at the IETF. This draft is on its way to being a published standard and is already implemented in all major TLS libraries.
- We participated in the design of next version (1.3) of the TLS protocol. We hosted an interim TLS working group meeting in Paris. Our proposals such as the session hash construction are now an integral part of the new design, and we continue consulting on the design and implementation of TLS.

6.2. Verification of Security Protocols in the Symbolic Model

Participants: Bruno Blanchet, Miriam Paiola, Robert Künnemann.

Miriam Paiola wrote and defended her PhD thesis on the verification of security protocols with lists [45].

Robert Künnemann published a paper at the IEEE S&P conference on how to extend symbolic cryptographic protocol verifiers to account for global state [60].

Bruno Blanchet published a tutorial on the protocol verifier ProVerif [66], as a follow-up to his teaching in the FOSAD'13 summer school last year.

The applied pi calculus is a widely used language for modeling security protocols, including as a theoretical basis of ProVerif. However, the seminal paper that describes this language (Abadi and Fournet, POPL'01) does not come with proofs, and detailed proofs for the results in this paper were never published. This year, Martin Abadi, Bruno Blanchet, and Cedric Fournet wrote detailed proofs for the main theorems of this paper. This work was also an opportunity to fix a few minor details in the results and to tune the calculus to improve it and make it closer to the input calculus of ProVerif. We plan to submit this work as a journal paper.

6.3. Verification of Security Protocols in the Computational model

Participants: Bruno Blanchet, David Cadé.

We worked on our computationally-sound protocol verifier CryptoVerif in two directions.

First, this verifier includes a specialized compiler that generates secure implementations of protocols from CryptoVerif specifications. We completed a journal version of the proof that this compiler preserves security, which is to appear in the Journal of Computer Security [48]

Second, Bruno Blanchet extended CryptoVerif with support for equational theories: associativity, commutativity, non-commutative and commutative groups, exclusive or. The goal is to be able to verify protocols that rely on the algebraic properties of groups and exclusive or. The extended tool is available at http://cryptoverif.inria.fr.

6.4. Computationally Complete Symbolic Attacker Models

Participants: Gergei Bana, Hubert Comon-Lundh.

A new approach to computational verification is to define a *computationally complete* symbolic attacker, so that a symbolic proof against this attacker can be shown to imply a computational proof of security. Following this line of inquiry, Gergei Bana and Hubert Comon-Lundh recently published work on proving computational reachability properties using symbolic techniques.

Gergei Bana (along with Hubert Comon-Lundh) published a paper on how to extend this work to prove stronger security properties expressed as equivalences [50]. Hence, the proof techniq can now be used also for properties like anonymity, strong secrecy etc. Besides being able to prove such properties, another advantage of this extension is that modern security properties of cryptographic primitives are also formulated in terms of indistinguishability, which makes it easier to translate the security properties cryptographers define to our language than before.

Using the computationally complete symbolic attacker, writing up a full, computationally sound proof (and identifying new attacks) for the NSL protocol when agents can run both roles, including running sessions with themselves. The proof is first attempted without any assumption other than that the encryption is CCA2 and that honest names are assigned at the beginning (that is, absolutely nothing about parsing: triples may be independent from pairs, pairing the projection of pairs may not give back the original item etc.). Along the way, we identified new attacks absent of some necessary parsing properties that implementations may not satisfy in general. Then with these additional parsing properties added to the properties satisfied by the implementation, we verified the protocol, namely secrecy, authentication and agreement. The project included graphical representation of the proof steps and the attacks. Type-flaw attacks that can be found in the literature have been reproduced this way, but a number of other attacks have also be revealed that cannot be found with the Dolev-Yao technique, and have not been found by other computational techniques either, although they are realistic. This is joint work with Pedro Adao of IST Lisbon. We hope to publish parts of this work to illustrate proving strategies. The current state of the writeup is available at http://prosecco.gforge.inria.fr/personal/gebana/nsl-long-both-roles.pdf

6.5. Authentication Attacks against Transport Layer Security

Participants: Karthikeyan Bhargavan [correspondant], Antoine Delignat-Lavaud, Cedric Fournet [Microsoft Research], Markulf Kohlweiss [Microsoft Research], Alfredo Pironti, Pierre-Yves Strub [IMDEA].

We discovered an important client impersonation attack on the Transport Layer Security protocol called the *triple handshake attack*. The attack is on the standard protocol and hence all compliant implementations were potentially at risk. Hence, we systematically followed responsible disclosure by notifying all major web browsers and TLS implementors, and then working with the TLS working group to design a countermeasure. The research results of this work were published at IEEE S&P [53].

To TLS implementors, we proposed short-term countermeasures that mitigated our attack, leading to security updates to all major web browsers: Google Chrome (CVE-2013-6628), Mozilla Firefox (CVE-2014-1491), Internet Explorer (CVE-2014-1771), Apple Safari (CVE-2014-1295), as well as to non-browser TLS libraries such as Oracle JSSE (CVE-2014-6457) and RSA BSAFE (CVE-2014-4630). For more details, see http://secure-resumption.com

To the TLS working group, we proposed a new cryptographic construction called the *session hash* that fundamentally alters the cryptographic core of TLS. This construction has now been adopted as a protocol extension to TLS 1.2 and has been integrated into the upcoming TLS 1.3. We expect an IETF standard for this construction to be published in early 2015.

While the triple handshake attacks primarily affect client-authentication, server authentication in HTTPS (HTTP over TLS) primarily relies on X.509 public key certificates. Antoine Delignat-Lavaud along with co-authors at Microsoft research published a paper at NDSS 2015 on a large-scale study of the Web PKI: how certificates are issued and used on the web [56]. Our work uncovered many unsafe practices and suggested best practices and new security policies.

Antoine Delignat-Lavaud also showed how the unsafe sharing of certificates across multiple websites could be exploited to fully compromise the same origin policty for websites, using an vulnerability called virtual host confusion. These results were discussed in a talk at BlackHat USA: for details see http://bh.ht.vc. A research paper on these attacks is forthcoming at WWW'2015.

6.6. A Verified Reference Implementation of Transport Layer Security

Participants: Benjamin Beurdouche [correspondant], Karthikeyan Bhargavan [correspondant], Antoine Delignat-Lavaud, Cedric Fournet [Microsoft Research], Markulf Kohlweiss [Microsoft Research], Alfredo Pironti, Pierre-Yves Strub [IMDEA], Santiago Zanella-Béguelin [Microsoft Research], Jean Karim Zinzindohoue.

Following on from previous work in the miTLS project, we published new versions of miTLS (http://mitls.org) that implemented various protocol extensions including the new session hash extension.

At CRYPTO 2014 [55], we published the first detailed cryptographic proof of an implementation of the TLS Handshake. The implementation consists of about 5000 lines of code and is equipped with about 2500 lines of security annotations written in F7, and a 3000 line EasyCrypt proof.

Currently, we are extending and improving this verified implementation to cover commonly used TLS extensions as well as TLS 1.3, the new version of TLS that we are actively involved in designing. We recently hosted a meeting of the TLS working group at Inria in Paris and are active members of the core working group.

In parallel, we have been analyzing other implementations of TLS and testing them against our implementation, both to ensure interoperability and to uncover bugs. Our analyses have led to the discovery of serious state machine vulnerabilities in many TLS implementations including Oracle JSSE, NSS, OpenSSL, SecureTransport, CyaSSL, Mono, and RSA BSAFE. On our recommendations, all these TLS libraries have issued important security updates in 2014.

6.7. Verified implementations of cryptographic primitives

Participants: Evmorfia-Iro Bartzia, Jean Karim Zinzindohoue, Pierre-Yves Strub, Karthikeyan Bhargavan.

Cryptographic libraries underpin the security of all security protocol implementations. A bug in the implementation of one primitive could enable an attacker to break the security of the full protocol. Hence, establishing the formal correctness of an efficient cryptographic mechanism is a much-desired but still open goal. We are investigating two directions of research towards this goal, specifically in the context of elliptic curve libraries.

Evmorfia-Iro Bartzia and Pierre-Yves Strub are building a Coq library that enables the precise proof of elliptic curve algorithms, and the automatic extraction of verified OCaml code that implements these algorithms. Their most recent result is the formal proof of a non-trivial theorem by Picard: the existence of an isomorphism between an elliptic curve and its Picard group of divisors. This work led to the publication "A formal library for Elliptic Curves in the Coq proof Assistant" and was presented at the ITP 2014 conference [51]. We have also been working on a formal proof of correctness of the GLV algorithm for scalar multiplication in Coq, using the above development and the CoqEal methodology. At present, we have an implementation of the algorithm in the OCaml language and a formal development regarding multiexponentiation, endomorphisms, scalar decomposition and coordinates in both affine and projective spaces. This work is still in progress.

Jean Karim Zinzindohoue and Karthikeyan Bhargavan are investigating the direct verification of implementations of the Curve25519 elliptic curve that is emerging as the preferred new curve for a variety of cryptographic standards, including TLS and the W3C web cryptography API. We use standard program verification tools such as the Frama-C/Why3 verification toolkit for a C implementation of Curve25519 and the F* typechecker for an OCaml implementation of the curve. This work is still in progress.

6.8. Dynamic Security Verification and Testing

Participants: Catalin Hritcu, Arthur Azevedo de Amorim, Zoi Paraskevopoulou, Nikolaos Giannarakis.

We investigated two directions in the runtime security verification of software and hardware systems.

Catalin Hritcu, Arthur Azevedo de Amorim, Nick Giannarakis, and their collaborators at University of Pennsylvania and Portland State University published work on *micro-policies* a generic framework for defining tag-based reference monitors on a simple tagged RISC processor. The framework was formalized and verified in the Coq proof assistant and was used to define and verify micro-policies for dynamic sealing, control-flow integrity, compartmentalization, and memory safety. This work resulted in publications at POPL 2014 [63], ASPLOS 2015 [58], and another paper is in submission.

Catalin Hritcu along with his co-authors worked on a testing framework for security and functional correctness. We published a journal paper about testing noninterference [68] and submitted an ANR JCJC grant pre-proposal on the whole project. Catalin Hritcu also worked with an intern Zoe Paraskevopoulou on this topic, who successfully defended her thesis at NTU Athens. We plan to publish a polished version of that in the near future.

6.9. Verified Security for Web Applications

Participants: Karthikeyan Bhargavan [correspondant], Chetan Bansal [Microsoft], Antoine Delignat-Lavaud, Sergio Maffeis [Imperial College London].

Karthikeyan Bhargavan, Antoine Delignat-Lavaud, and co-authors published a tutorial on Defensive JavaScript, a typed subset of JavaScript that is designed to be used for security-critical components such as cryptographic libraries that may be deployed within untrusted web pages. This tutorial was published as a follow-up of Karthikeyan Bhargavan's lectures at the FOSAD'13 summer school [65].

Karthikeyan Bhargavan, Antoine Delignat-Lavaud, and co-authors also published a journal version of their work on the WebSpi web security modeling library [47], one of the few formal models that captures the detailed security assumptions of various web mechanisms.

Karthikeyan Bhargavan along with collaborators at Microsoft Research published a paper at POPL 2014 on TS*: a new gradual type system for a large subset of JavaScript [47]. We showed how to compile and safely deploy well-typed TS* programs as standard JavaScript in websites. Such programs preserve their types even if other code running on the website is malicious. Our work was used as a basis for further work on the TypeScript compiler and typechecker developed at Microsoft.

6.10. Electronic Voting and Auctions

Participants: Benjamin Smyth [correspondant], Elizabeth Quaglia, Adam Mccarthy, David Bernhard.

Benjamin Smyth continued his work on proving privacy properties of electronic voting protocols. Smyth and Bernhard worked on a new formal definition of ballot secrecy that works even if the bulletin board (used to publish votes) is malicious [69].

Benjamin Smyth, Elizabeth Quaglia, and Adam McCarthy observed that existing electronic voting schemes could be used as core building blocks for electronic auction protocols. Using this link, they build two new e-auction protocols Hawk and Aucitas by building on top of the e-voting protocols Helios and Civitas resp. They prove that their protocols enjoy many desired security properties. This result was published at Financial Cryptography 2014 [61].

QUANTIC Team

5. New Results

5.1. Highlights of the Year

- Experimental results in continuous measurement of error syndromes for a quantum error correction scheme developed by Mazyar Mirrahimi and his former PhD student Zaki Leghtas in close collaboration with the teams of Michel Devoret and Robert Schoelkopf (Department of Applied Physics of Yale University) have been published in Nature [13].
- Theoretical proposal on a new paradigm for universal quantum computation [12] has been chosen by the editors of the New Journal of Physics as an IOPselect paper for the novelty, significance and potential impact on future research.
- The EPOQ2 ANR Young Researcher project, led by Mazyar Mirrahimi, was highlighted in the 2013 annual report of Agence Nationale de la Recherche.

5.2. Dynamically protected cat-qubits: a new paradigm for universal quantum computation

Participant: Mazyar Mirrahimi.

In a close collaboration with the teams of Michel Devoret, Robert Schoelkopf and Liang Jiang (Department of Applied Physics, Yale university) and in particular a former member of our group, Zaki Leghtas, we have presented a new hardware-efficient paradigm for universal quantum computation. This paradigm is based on encoding, protecting and manipulating quantum information in a quantum harmonic oscillator. This proposal exploits multi-photon driven dissipative processes to encode quantum information in logical bases composed of Schrödinger cat states. More precisely, we consider two schemes. In a first scheme, a two-photon driven dissipative process is used to stabilize a logical qubit basis of two-component Schrödinger cat states. While such a scheme ensures a protection of the logical qubit against the photon dephasing errors, the prominent error channel of single-photon loss induces bit-flip type errors that cannot be corrected. Therefore, we have considered a second scheme based on a four-photon driven dissipative process which leads to the choice of four-component Schrödinger cat states as the logical qubit. Such a logical qubit can be protected against single-photon loss by continuous photon number parity measurements. Next, applying some specific Hamiltonians, we have provided a set of universal quantum gates on the encoded qubits of each of the two schemes. In particular, we have illustrated how these operations can be rendered fault-tolerant with respect to various decoherence channels of participating quantum systems. Finally, we have also proposed experimental schemes based on quantum superconducting circuits and inspired by methods used in Josephson parametric amplification, which should allow to achieve these driven dissipative processes along with the Hamiltonians ensuring the universal operations in an efficient manner.

This proposal was published in New Journal of Physics [12] and has also been chosen by the editor as an IOPselect paper for the novelty, significance and potential impact on future research.

5.3. Tracking photon jumps with repeated quantum non-demolition parity measurements

Participant: Mazyar Mirrahimi.

Quantum error correction (QEC) is required for a practical quantum computer because of the fragile nature of quantum information. In quantum error correction, information is redundantly stored in a large quantum state space and one or more observables must be monitored to reveal the occurrence of an error, without disturbing the information encoded in an unknown quantum state. Such observables, typically multi-quantum-bit parities, must correspond to a special symmetry property inherent in the encoding scheme. Measurements of these observables, or error syndromes, must also be performed in a quantum non-demolition way (projecting without further perturbing the state) and more quickly than errors occur. Previously, quantum non-demolition measurements of quantum jumps between states of well-defined energy have been performed in systems such as trapped ions, electrons, cavity quantum electrodynamics, nitrogen?vacancy centres and superconducting quantum bits. So far, however, no fast and repeated monitoring of an error syndrome had been achieved. Mazyar Mirrahimi has participated to an experiment performed by the group of Robert Schoelkopf (Department of Applied Physics, Yale University) where the quantum jumps of a possible error syndrome, namely the photon number parity of a microwave cavity, were tracked by mapping this property onto an ancilla quantum bit, whose only role is to facilitate quantum state manipulation and measurement. This quantity is just the error syndrome required in a QEC scheme proposed by Mazyar Mirrahimi and his former PhD student, Zaki Leghtas, and in a close collaboration with the teams of Michel Devoret and Robert Schoelkopf. This scheme should lead to a hardware-efficient protected quantum memory using Schrödinger cat states (quantum superpositions of different coherent states of light) in a harmonic oscillator [4]. We demonstrated the projective nature of this measurement onto a region of state space with well-defined parity by observing the collapse of a coherent state onto even or odd cat states. The measurement is fast compared with the cavity lifetime, has a high single-shot fidelity and has a 99.8 per cent probability per single measurement of leaving the parity unchanged. In combination with the deterministic encoding of quantum information in cat states realized earlier [10], the quantum non-demolition parity tracking that we have demonstrated represents an important step towards implementing an active system that extends the lifetime of a quantum bit. This result was published in Nature [9].

5.4. Dissipation-induced continuous quantum error correction for superconducting circuits

Participants: Joachim Cohen, Mazyar Mirrahimi.

Quantum error correction (QEC) is a crucial step towards long coherence times required for efficient quantum information processing (QIP). One major challenge in this direction concerns the fast real-time analysis of error syndrome measurements and the associated feedback control. Recent proposals on autonomous QEC (AQEC) have opened new perspectives to overcome this difficulty. As a sequel to our recent contributions to autonomous stabilization of maximally entangled states of superconducting qubits [53],[8], we have designed an AQEC scheme based on quantum reservoir engineering adapted to superconducting qubits. We have focused on a three-qubit bit-flip code, where three transmon qubits are dispersively coupled to a few low-Q resonator modes. By applying only continuous-wave drives of fixed but well-chosen frequencies and amplitudes, we engineer an effective interaction Hamiltonian to evacuate the entropy created by eventual bit-flip errors. We have provided a full analytical and numerical study of the protocol, while introducing the main limitations on the achievable error correction rates. This result was published in Physical Review A [11].

5.5. Continuous generation and stabilization of mesoscopic field superposition states in a quantum circuit

Participants: Ananda Roy, Mazyar Mirrahimi.

While dissipation is widely considered as being harmful for quantum coherence, it can, when properly engineered, lead to the stabilization of non-trivial pure quantum states. In a close collaboration with the teams of Michel Devoret and Douglas Stone (Department of Applied Physics, Yale University), and in the framework of a 6 months visit by Ananda Roy (PhD student at Yale), we proposed a scheme for continuous generation and stabilization of Schrödinger cat states in a cavity using dissipation engineering [15]. The scheme consists in first generating non-classical photon states with definite parity by means of a two-photon drive and dissipation, and then stabilizing these transient states against single-photon decay. The single-photon stabilization is autonomous, and is implemented through a second engineered bath, which exploits the photon number dependent frequency-splitting due to Kerr interactions in the strongly dispersive regime of circuit QED. Starting with the Hamiltonian of the baths plus cavity, we derived an effective model of only the cavity photon states along with analytic expressions for relevant physical quantities, such as the stabilization rate. The deterministic generation of such cat states is one of the key ingredients in performing universal quantum computation.

5.6. Extending robustness and randomization from consensus to symmetrization algorithms

Participant: Alain Sarlette.

In the framework of a collaboration with Francesco Ticozzi (University of Padova) on common points between quantum and classical network dynamics, we developed a general "symmetrization" framework which covers robust ways to generate dynamics in several algorithmic and control contexts [18]. The starting point was the question of generalizing so-called "consensus" algorithms to networks composed of quantum units. In order to define state information exchange without requiring state communication (an impossible feat given the quantum no-cloning theorem), an operational viewpoint on consensus had been proposed by Alain Sarlette and co-authors in the previous year. In this new result, the scope of this operational viewpoint is considerably extended by considering it as a "symmetrization" procedure with respect to some discrete group, completely abstracting away the actual action space. It is shown that this abstraction covers existing procedures ranging from network synchronization to random state generation (not in networks) and averaging-based open-loop control procedures. The interest of viewing those procedures under the common "symmetrization" framework proposed is twofold: convergence proofs follow from a general result that we have established; and robustness to randomized actions and (specific) parameter uncertainties is shown to carry over from the "consensus" literature. It is further anticipated that the approach might be a guideline for new algorithmic designs in the future.

5.7. Accelerating consensus by spectral clustering and polynomial filters

Participant: Alain Sarlette.

The previous work of Alain Sarlette about quantum consensus and symmetrization has been further explored towards quantum-induced accelerations of algorithms, thermalization processes and random walks. This work is still at a preliminary stage. It has been noticed that some non-quantum acceleration possibilities were not fully explored and this has led to two publications that establish preliminary clarifications for our main goal. In [17], a standing conjecture has been proved which claims that if only the spectral gap of a graph is known (i.e. a bound on its lowest and largest eigenvalues), then by adding m local memories to each node no faster convergence can be obtained than by adding $m = 2$ local memories. The conjecture is proved with an analogy to root locus techniques, and a network-centric (e.g. information-theory-based) argument for this fact is currently missing, but at least the fact has been established. This allows for direct comparisons with "quantum random walk" accelerations, which obtain the same speed as $m = 2$ but with a different tweak, that is based among others on more knowledge of the network structure. In this spirit, we have clarified in [16] how classical consensus with time-varying filters can benefit from knowledge of extra bounds on the graph eigenvalue locations (without knowing them exactly, which is the case considered in the existing literature). This work also observes how the speed-up trades off with robustness to network modifications.

5.8. Integral control on Lie groups

Participant: Alain Sarlette.

A big challenge for the long-term control of interacting networks is their robustness to systematic biases. Integral control is a standard way to counter them when a target output can be measured. This method has been originally proposed, and extensively studied, for linear systems. However when the system (output) evolves on a nonlinear state space, the standard "integration" technique cannot be straightforwardly applied. Especially for global motions on spaces like the circle, sphere or (real or complex) rotation groups, the output integration viewpoint becomes problematic. We have hence proposed a new viewpoint on integral control, based on integrating the intended input [19]. For linear state spaces, it is equivalent to the standard definition. For nonlinear state spaces, this viewpoint can be transposed verbatim modulo introduction of a transport map on the tangent bundle, which is almost always present for control design purposes. In particular for systems on Lie groups, which are ubiquitous in robotics and in quantum physics, a full analysis of fully actuated systems has been proposed. The more challenging extension to underactuated systems is underway.

RAP Project-Team

4. New Results

4.1. Random Graphs

Participants: Nicolas Broutin, Henning Sulzbach.

4.1.1. Universality of scaling limits of random graphs

Random graphs are one of the most studied models of networks, and they turn out to be related to crucial questions in physics about the behaviour of matter at the phase transition, or in combinatorial optimization about the hardness of computation. In recent years, we have constructed the scaling limit of the classical Erdos-Renyi random graph model, and conjectured that this limit also happened to be universal.

The funding of the Associated Team RNA has permitted to invite Shankar Bhamidi. During his visit, we have worked and found a new way to construct the scaling limit of random graph processes in the critical window. This method is especially important since it is robust enough to prove universality of the limit, that is that many models have the same limit. The method relies on the dynamics of the coalescence of clusters as the edges are added, and allows us to hope for proofs that would be able to treat the more complex geometric models.

4.1.2. Cutting down random tree and the genealogy of fragmentations

The study of the internal structure of random combinatorial object such as graphs and trees led to question about whether such objects exhibit invariance by certain complex surgical operations (disconnect some pieces, and re-attach them somewhere else). In the context of graphs, this is related to the so-called self-organized criticality: certain distributions that yield fractal objects should naturally appear in nature because they are the fixed points of some recombination procedures. In the context of trees, it turns out that certain fragmentations arising when chopping off a random tree have a genealogy that has the same distribution as the original tree. We have investigated this with Minmin Wang, and obtained results about p-trees and the genealogy of the fragmentation on Aldous' celebrated continuum random tree. These may also be interpreted in terms of complex path transformations for Brownian excursions and other random processes with exchangeable increments, and hence relate to very classical questions in probability theory.

4.1.3. New encodings for combinatorial coalescent processes

In 2013, we had constructed the scaling limit of the minimum spanning tree of a complete graph using crucial information about the scaling limit of random graphs, and especially about the way the cluster merge as the edges are added in the graph. With J.-F. Marckert (LaBRI, Bordeaux) we have found a novel construction of the important multiplicative coalescent that describes how the connected components of a random graph coalesce as the edges are added. This unveils yet more interesting links between the minimum spanning tree and the random graph, since Prim's celebrated algorithm is used to construct a consistent ordering of the vertices that ensures that the connected components are intervals.

4.1.4. Navigation in random Delaunay triangulations

Navigation or routing algorithms are fundamental routines: in order to solve many problems, one of the first steps consists in locating a node in a data structure. Unfortunately, the current algorithms are based on heuristics and very few rigorous results about the performance of such algorithms are known when the model for data is more realistic than the worst-case.

With O. Devillers and R. Hemsley, we have initiated a program that aims at finding rigourous estimates for the performance of routing algorithms in geometric structures such as Delaunay tesselations. So far we have managed to develop some tools that permitted us to analyse a simple algorithm. Although this algorithm has been designed for most of the analysis to work, this work paves the way towards the rigorous analysis of other more natural and widely used algorithms.

4.1.5. *Connectivity and sparsification of sparse wireless networks*

Many models of wireless networks happen to be connected only when the average degree is tending to infinity with the size of the network, more precisely when it is about the logarithm of the number of nodes. This raises questions about the potential issues in scaling such models. With L. Devroye (McGill) and G. Lugosi (ICREA and Pompeu Fabra), we have worked at analysing models in which we try to construct connected or almost connected networks in a distributed way (that is that no global optimization is allowed in designing the network, and every device should proceed in the same way to choose its neighbors). We have managed to analyse an algorithm for constructing such a network, and to obtain tight results about the number of links that a typical device should have in order for the global network to be connected. We further proved that this is asymptotically optimal when one only requires that most nodes should be in the same connected component.

4.2. Resource Allocation Algorithms in Large Distributed Systems

Participants: Christine Fricker, Philippe Robert, Guilherme Thompson.

This is a collaboration with Fabrice Guillemin from Orange Labs which started in February 2014.

4.2.1. *Controlling impatience in cellular networks using QoE-aware radio resource allocation*

Impatience of users when using a data service has a major impact on the quality of service offered by telecommunication networks, especially in cellular networks with scarce radio resources. Impatience is negative for users, it is due to many factors related to the performance of servers, customer devices, etc., but also to bandwidth sharing in the network.

While impatience can be seen as a negative phenomenon, it can also be used as a lever to discourage customers when the system becomes too much overloaded. This can be achieved in cellular networks by modulating the capacity available to customers being at a certain distance of the antenna. This general idea can be applied in several manners and can be viewed as a network optimization mechanism. In this paper, we reuse the general framework of α-fair scheduler in order to perform this control. This has the advantage of being easy to implement in realistic settings as α-fair schedulers (and especially the Proportional Fair (PF) one) are widely adopted in mobile networks. This also reduces the dimension of our problem as it narrows the optimization problem to the tuning of a single parameter α.

In order to achieve this goal, we first derive a model for reneging probabilities under a general α-fair scheduler. In particular, we consider a heavy load regime and develop a fluid flow analysis of impatience in cellular networks. We notably establish a fixed point formulation for the computation of the reneging probability and introduce a new metric, namely QoE perturbation, expressing how much a particular flow impacts the reneging probability in the system. We then use this QoE perturbation metric to design of a new radio resource management scheme that controls the parameter of the scheduler in order to reduce the global reneging in the system. For instance, recognizing that customers far from the base station degrade the global performance of the system, impatience and α-fair scheduling can be used to discourage those customers and in some sense to perform an implicit admission control in order to optimize the use of radio resources.

4.2.2. *Resource Allocation in Large Data Centers*

The goal of this study is to investigate the design of allocation algorithms of requests requiring different classes of quality of video streams as well as their performances. The class of algorithms considered may downgrade the quality of some of the transmission to maximize the utilization of the servers.

4.3. Stochastic networks: large bike sharing systems

Participants: Christine Fricker, Hanène Mohamed, Cédric Bourdais, Yousra Chabchoub.

Vehicle sharing systems are becoming an urban mode of transportation, and launched in many cities, as Velib' and Autolib' in Paris. One of the major issues is the availability of the resources: vehicles or free slots to return them. These systems became a hot topic in Operation Research and now the impact of stochasticity on the system behavior is commonly admitted. The problem is to understand their behavior and how to manage them in order to provide both resources to users.

Our stochastic model is the first studying the impact of the finite number of spots at the stations on the system behavior.

With Danielle Tibi, we use limit local theorems to obtain the asymptotic stationary joint distributions of several node (station or route) states when the system is large (both numbers of stations and bikes), also in the case of finite capacities of the stations. This gives an asymptotic independence preperty for node states. This widely extends the existing results on heterogeneous bike-sharing systems.

Second we investigate the impact of finite capacity of stations and reservation in car-sharing systems. The large-scale asymptotic joint stationary distribution of the numbers of vehicles and reserved parking places is given as the joint distribution in a tandem of queues with a constrained total capacity where rates are solutions of a system of two fixed point equations. Analytical expressions are given for performance in light and heavy traffic cases. As expected, reservation impact drastically increases with traffic. Even if the equilibrium is identified and analyzed, the question of convergence is still open.

JC Decaux provides us data describing Velib' user trips. These data are useful to measure the system behavior. With Yousra Chabchoub, we test clustering to obtain a typology of the stations. Then we focus on the resources availability (free docks and available bikes) and separate the Velib' stations into three clusters (balanced, overloaded and underloaded stations), using Kmeans clustering algorithm, along with the Dynamic Time Wraping (DTW) metric. We choose to update the centers of the clusters using the efficient Dtw Barycenter Averaging (DBA) method.

4.4. Scaling Methods

Participants: Philippe Robert, Wen Sun, Mohammadreza Aghajani.

4.4.1. Fluid Limits in Wireless Networks

This is a collaboration with Amandine Veber (CMAP, École Polytechnique). The goal is to investigate the stability properties of wireless networks when the bandwidth allocated to a node is proportional to a function of its backlog: if a node of this network has x requests to transmit, then it receives a fraction of the capacity proportional to $\log(1+x)$, the logarithm of its current load. A fluid scaling analysis of such a network is presented. We have shown that the interaction of several time scales plays an important role in the evolution of such a system, in particular its coordinates may live on very different time and space scales. As a consequence, the associated stochastic processes turn out to have unusual scaling behaviors which give an interesting fairness property to this class of algorithms. A heavy traffic limit theorem for the invariant distribution has also been proved. A generalization to the resource sharing algorithm for which the $\log$ function is replaced by an increasing function. This year we completed the analysis of a star network topology with multiple nodes. Several scalings were used to describe the fluid limit behaviour.

4.4.2. The Time Scales of a Transient Network

The Distributed Hash Table (DHTs) consists of a large set of nodes connected through the Internet. Each file contained in the DHT is stored in a small subset of these nodes. Each node breaks down periodically and it is necessary to have back-up mechanisms in order to avoid data loss. A trade-off is necessary between the bandwidth and the memory used for this back-up mechanism and the data loss rate. Back-up mechanisms already exist and have been studied thanks to simulation. To our knowledge, no theoretical study exists on this topic. With a very simple centralized model, we have been able to emphasise a trade-off between capacity and life-time with respect to the duplication rate. From a mathematical point of view, we are currently studying different time scales of the system with an averaging phenomenon.

4.5. Stochastic Models of Biological Networks

Participants: Renaud Dessalles, Sarah Eugene, Emanuele Leoncini, Philippe Robert.

4.5.1. Stochastic Modelling of self-regulation in the protein production system of bacteria

This is a collaboration with Vincent Fromion from INRA Jouy-en-Josas, which started on December 2014.

In procaryots cells (e.g. E. Coli. or B. Subtilis) the protein production system has to produce in a cell cycle (i.e. less than one hour) more than 10^6 molecules of more than 2500 kinds, each having different level of expression. The bacteria uses more than 85% of its resources to the protein production. Gene expression is a highly stochastic process: bacteria sharing the same genome, in a same environment will not produce exactly the same amount of a given protein. Some of this stochasticity can be due to the system of production itself: molecules that take part in the production process move freely into the cytoplasm and therefore reach any target in the cell after some random time; some of them are present in so much limited amount that none of them can be available for a certain time; the gene can be deactivated by repressors for a certain time etc...

We study the integration of several mechanisms of regulation and their performances in terms of variance and distribution. All molecules are supposed to move freely into the cytoplasm, it is assumed that the the encounter time between a given entity and its target is exponentially distributed.

4.5.1.1. Transcription-translation model for all proteins

The first model that has been studied integrates the production of all the proteins. Each gene has to be transcribed in mRNA and each mRNA has to be translated in protein. The transcription step needs a RNA-Polymerase molecule that is sequestered during the time of elongation. Likewise, each mRNA needs a ribosome in order to produce a protein. RNA-Polymerases/Ribosomes are present in limited amount and the genes/mRNAs sequester these molecules during the whole the time of elongation. Finally each mRNA has an exponentially distributed lifetime with an average value of 4 min and the proteins disappear at a rate of one hour, hence simulating the global dilution in the growing bacteria.

This global sharing of Ribosomes/RNA-Polymerases among all proteins induces a general regulation: each gene competing to each other to have access to these common resources. Because of the parameters of affinity (between gene and RNA-Poymerase and between mRNA and ribosome) are specific to each gene, it allows a large range of average protein production but induce some noise, especially for highly expressed proteins.

We developed a Python simulation, and using the biological experiments of Tanichuchi et al. (2010), and we have investigated a biologically coherent range of parameters. By making the simulations, we have been able to reproduce certain aspects of the biological measures, especially for the high amount of noise for well expressed proteins.

4.5.1.2. Simple feedback model

We have also investigated the production of a single protein, with the transcription and the translation steps, but we also introduced a direct feedback on it: the protein tends to bind on the promoter of its own gene, blocking therefore the transcription. The protein remains on it during an exponential time until its detachment caused by thermal agitation.

The mathematical analysis aims at understanding the nature of the internal noise of the system and to quantify it. We try to determine if, for instance, for the same average protein level, the feedback permits a noise reduction of protein distribution compared to the "open loop" model; or if it rather allows a better efficiency in case of a change of command for a new level of production (due, for example, to a radical change in the environment) by reducing the respond time to reach this new average.

4.5.2. Stochastic Modelling of Protein Polymerization

This is a collaboration with Marie Doumic, Inria MAMBA team.

Our work focuses on the study of the polymerization of protein. This phenomenon is involved in many neurodegenerative diseases such as Alzheimer's and Prion diseases, e.g mad cow. In this context, it consists in the abnormal aggregation of proteins. Curves obtained by measuring the quantity of polymers formed in in vitro experiments are sigmoids: a long lag phase with almost no polymers followed by a fast consumption of all monomers. Furthermore, repeating the experiment under the same initial conditions leads to somewhat identical curves up to a translation.

The first study we did proposed a simplified stochastic model to analyze this phenomenon. For this model, when the volume gets large, the quantity of polymers has the typical sigmoidal shape. A second order result has also been obtained for this model. We were able to compute the asymptotic distribution of the lag time and express its variance. The parameters of the model have been obtained by using data given by Wei-Feng Xue, University of Kent.

The current project concerns a more sophisticated mathematical model. Indeed, we have added a conformation step: before polymerizing, proteins have to misfold. This step is very quick and remains at equilibrium during the whole process. Nevertheless, this equilibrium depends on the polymerization which follows the conformation step: this modelling leads to the study of averaging principles.

REALOPT Project-Team

6. New Results

6.1. Highlights of the Year

- Olivier Beaumont and Lionel Eyraud-Dubois have received the HiPC best paper award for their work on resource allocation for large scale virtualized platforms with reliability guarantees. They provided a formulation based on a thorough analysis of a real life usage trace, and a very efficient two-step allocation algorithm.
- The team organized the annual conference of the French Operations Research Society ROADEF14 in Feb 2014.
- An Inria Innovation Lab has been created between Realopt and Ertus Consulting.
- The SAMBA associated team project with Brazil was renewed for 3 years including new collaborators from Chili.
- François Vanderbeck was invited as a plenary speaker at the conference OPTIMIZATION 2014, in Portugal [19].

6.2. Automation and combination of linear-programming based stabilization techniques in column generation

We reviewed in [88] stabilization techniques that can improve in practice the convergence of a column generation algorithm. Proximal methods based on penalising the deviation from the incumbent dual solution have become standards of the domain. However, the analysis of such methods is important to understand the mechanism on which they rely, to appreciate the difference between methods, and to derive intelligent schemes to adjust their parameters. As stabilization procedures for column generation can be viewed as cutting plane strategies in the dual problem, the link with cutting plane separation strategies can be exploited to enlarge the scope of methods and to refine their analysis. In [24], [40], we focus on stabilization schemes that rely solely on a linear programming (LP) solver for the column generation master program. This restrictive scope captures the most common implementations where one uses an LP solver to handle the master program. For dual price smoothing techniques, we analyse the link with the in-out separation strategy and we derive generic convergence properties. For penalty function methods as well as for smoothing, we describe proposals for parameter self-adjusting schemes. Such schemes make initial parameter tuning less of an issue as corrections are made. Also, the dynamic adjustments, compared to a static setting, allows to adapt the parameters to the phase of the algorithm. We provide extensive test reports that highlight the comparative performances of such scheme and validate our self-adjusting parameter scheme. Furthermore, our results show that using smoothing in combination with penalty function yields a cumulative effect on convergence speed-ups [35]. We have also consider other stabilization strategies inspired form algorithmic strategies have been designed to accelerate convergence of cutting plane algorithms in mixed integer programming. In [37], we show that the "Multi-Point Separation" strategy translates into a column generation stabilization technique that consists in restricting the dual solution to be in the convex hull of the selected multi-point set. We have also considered other stabilization strategies inspired from algorithmic strategies that have been designed to accelerate convergence of cutting plane algorithms in mixed integer programming. In [37], we show that the "Multi-Point Separation" strategy translates into a column generation stabilization technique that consists in restricting the dual solution to be in the convex hull of the selected multi-point set.

6.3. Multi-Stage Column generation strategies

In [39], we propose another mechanism to improve the performance of column generation algorithms. We study the application of branch-and-price approaches to the automatic version of the Software Clustering Problem. To tackle this problem, we apply the Dantzig-Wolfe decomposition to a formulation from literature. Given this, we present two Column Generation (CG) approaches to solve the linear programming relaxation of the resulting reformulation: the standard CG approach, and a new approach, which we call Staged Column Generation (SCG). Also, we propose a modification to the pricing subproblem that allows to add multiple columns at each iteration of the CG. We test our algorithms in a set of 45 instances from the literature. The proposed approaches were able to improve the literature results solving all these instances to optimality. Furthermore, the SCG approach presented a considerable performance improvement regarding computational time, number of iterations and generated columns when compared with the standard CG as the size of the instances grows.

6.4. Aggregation techniques to reduce the size of column generation models

We proposed an aggregation method to reduce the size of column generation (CG) models for a class of set-covering problems in which the feasible subsets depend on a resource constraint. The aggregation relies on a correlation between the resource consumption of the elements and the corresponding optimal dual values. The aggregated model obtained allows to find good quality lower bounds more rapidly than the original CG algorithm. The speedup is due to less primal and dual variables in the master, and to an aggregated pricing sub-problem. To guarantee optimaly, we designed an algorithm that iteratively refines the aggregation until the CG optimum is reached. Computational results prove the usefulness of our methods.

6.5. Dual-feasible functions

Dual-feasible functions have been used in the past to compute fast lower bounds and valid inequalities for different combinatorial optimization and integer programming problems. Until now, all the dual-feasible functions proposed in the literature were 1-dimensional functions, and were defined only for positive arguments. In [12] we extended the principles of dual-feasible functions to the m-dimensional case by introducing the concept of vector packing dual-feasible function. We explored the theoretical properties of these functions in depth, and we proposed general schemes for generating some instances of these functions. Additionally, we proposed and analyzed different new families of vector packing dual-feasible functions. All the proposed approaches were tested extensively using benchmark instances of the 2-dimensional vector packing problem. Our computational results showed that these functions can approximate very efficiently the best lower bounds for this problem. In a second paper, currently submitted to a journal, we show that extending these functions to negative arguments raises many issues. Additionally, we describe different construction principles to obtain dual-feasible functions with domain and range $\mathbb{R}$. Specific instances obtained from these principles are proposed and analyzed.

6.6. Resource Allocation and Scheduling in Large Scale Distributed Platforms.

We have considered several problems arising in the context of large scale platforms, that are characterized by their heterogeneity, the difficulty of predicting performance and the risk failures. In [13], we concentrate on heterogeneity issues in collective communication schemes where the goal is to broadcast a message to a set of nodes. In particular, we consider a realistic model in the context of large scale distributed platforms where some nodes may lie behind NATs or firewalls and may be therefore unable to forward the message between them. In [21], [20], we consider resource allocation problems that arise in large scale data centers. In [20], we analyze the main characteristics of the services in a huge trace corresponding to an actual data center and that has been released recently by google. In the same context, in [21], we concentrate on issues related to fault tolerance by over subscribing services in order to guarantee quality of service in a failure prone environment. At last, the difficulty to predict the actual performance of resources made it very popular to rely on dynamic scheduling algorithms where scheduling decisions are made at runtime. In [22], we analyze the performance of such a dynamic scheduling algorithm in terms on number of induced communications for outer product and matrix multiplication kernels.

6.7. Employee timetabling with time varying demand

We addressed a multi-activity tour scheduling problem with time varying demand. The objective is to compute a schedule for a fixed roster in order to minimize the over-coverage and the under-coverage of different parallel activity demands along a planning horizon. Numerous complicating constraints are present in our problem: all employees are different and can perform several different activities during the same day-shift, lunch breaks and pauses are flexible, demand is given for 15 minutes periods. To the best of our knowledge, the work in [29] is the first attempt to combine days-off scheduling, shift scheduling, shift assignment, activity assignment, pause and lunch break assignment. To solve this problem, we developed several methods: a compact linear Mixed Integer Programming model, a branch-and-price like approach with a nested dynamic program to solve heuristically the subproblems, a diving heuristic, and a greedy heuristic based on our subproblem solver. The computational results, based on both real cases and instances derived from real cases, demonstrate that our methods are able to provide good quality solutions in a short computing time. Our algorithms are now embedded in a commercial software, which is already in use in a mini-mart company.

6.8. Time-dependent formulations for routing problems

The paper [16] presents a new formulation for the Time-Dependent Travelling Salesman Problem (TDTSP). We start by reviewing well known natural formulations with some emphasis on the formulation by Picard and Queyranne (1978). The main feature of this formulation is that it uses, as a subproblem, an exact description of the n-circuit problem. Then, we present a new formulation that uses more variables and is based on using, for each node, a stronger subproblem, namely a n-circuit subproblem with the additional constraint that the corresponding node is not repeated in the circuit. Although the new model has more variables and constraints than the original PQ model, the results given from our computational experiments show that the linear programming relaxation of the new model gives, for many of the instances tested, gaps that are close to zero. Thus, the new model is worth investigating for solving TDTSP instances. We have also provided a complete characterization of the feasible set of the corresponding linear programming relaxation in the space of the variables of the PQ model. This characterization permits us to suggest alternative methods of using the proposed formulations.

A well-known formulation for the unit-demand capacitated vehicle routing problem uses a single commodity flow system to represent the delivery of the items. The vehicle capacity is modeled by imposing a maximum capacity on the arcs used by the flow. In [30], we used a time-dependent formulation for the problem to derive, by projection, tighter bounding inequalities on the arcs. The first experiments show that these new inequalities permit to improve significantly the linear relaxation bound of the single commodity flow formulation. We are currently studying separation algorithms in order to generate dynamically these new inequalities.

6.9. Vehicle routing for dial-a-ride problems

Static and deterministic vehicle routing problems cannot be used in many real-life systems, as input data are not reliable and revealed over time. In [11], we study a pickup and delivery problem with time windows accounting for maximum ride time constraints – the so-called dial-a-ride problem – in its static and dynamic variant, and we make specific proposal on robust optimization models for this problem. To solve the static model, we develop a branch-and-price approach that handles ride time constraints in the process of generating feasible vehicle routes in the course of the optimization procedure. The work is focussed on the pricing problem solver and acceleration techniques for the branch-and-price approach. Our numerical results show that the method is competitive compared to existing approaches that are based on branch-and-cut. In the dynamic context, where some input data are revealed or modified over time, we apply our branch-and-price algorithm for re-optimization in a rolling horizon approach.

6.10. A MILP approach to minimize the number of late jobs with and without machine availability constraints

The study in [14] investigates scheduling problems that occur when the weighted number of late jobs that are subject to deterministic machine availability constraints have to be minimized. These problems can be modeled as a more general job selection problem. Cases with resumable, non-resumable, and semi-resumable jobs as well as cases without availability constraints are investigated. The proposed efficient mixed integer linear programming approach includes possible improvements to the model, notably specialized lifted knapsack cover cuts. The method proves to be competitive compared with existing dedicated methods: numerical experiments on randomly generated instances show that all 350-job instances of the test bed are closed for the well-known problem $1|r_i|\sum w_i U_i$. For all investigated problem types, 98.4% of 500-job instances can be solved to optimality within one hour.

6.11. Two phase solution for an intelligent moving target search problem based on a 0–1 linear model

We developed a generic discrete model for the moving, intelligent target problem. Our objective is to maximise the probability of detection of the moving target with respect to target and searcher's constraints. The solution method proposed in [15] is composed of two stages. The first one aims at providing a large-scale strategy based on an Integer Linear Program approach. As a direct solution of this problem is not practically possible, we use a decomposition of the problem into a searcher's strategy on one side, and the target's strategy on the other side. A good strategy for the searcher is determined using a sliding window procedure. Concerning the target, our approach consists in simulating some of the target's possible strategies and considering each of these possibilities as an independent and deterministic entity. The second stage is dedicated to adjusting the large-scale strategy provided by stage 1. Numerical results are presented so as to assess the impact of our approach.

6.12. Computing the Chromatic index and clique numberof special graphs

In our paper [17] on the strong chromatic index of planar graphs with large girth, we prove that every planar graph with maximum degree Δ (let Δ be an integer) and girth at least $10\Delta + 46$ is strong $(2\Delta - 1)$-edge-colorable, that is best possible (in terms of number of colors) as soon as G contains two adjacent vertices of degree Δ. This improves the best previous result when $\Delta \geq 6$. In [18] we show how one can compute the clique number of a-perfect graphs in polynomial time. A main result of combinatorial optimization is that clique and chromatic number of a perfect graph are computable in polynomial time (Grötschel, Lovasz and Schrijver 1981). This result relies on polyhedral characterizations of perfect graphs involving the stable set polytope of the graph, a linear relaxation defined by clique constraints, and a semi-definite relaxation, the Theta-body of the graph. A natural question is whether the algorithmic results for perfect graphs can be extended to graph classes with similar polyhedral properties. In [18] we consider a superclass of perfect graphs, the a-perfect graphs, whose stable set polytope is given by constraints associated with generalized cliques. We show that for such graphs the clique number can be computed in polynomial time as well. The result strongly relies upon Fulkersons's antiblocking theory for polyhedra and Lovasz's Theta function.

REGAL Project-Team

5. New Results

5.1. Highlights of the Year

- *Garbage collection for big data on large-memory NUMA machines*. We developed NumaGiC, a high-throughput garbage collector for big-data algorithms running on large-memory NUMA machines (see Section 4.1). This result, a collaboration with the Whisper team, will be presented at ASPLOS 2015 [29].
- *Explicit consistency*. We propose an alternative approach to the strong-vs.-weak consistency conundrum, *explicit consistency*. Static analysis identifies precisely what is the minimal amount of synchronisation that is necessary to maintain the invariants required by an application (see Section 5.3.11). This result will be presented at EuroSys 2015 [53].
- *Lower bounds and optimality for CRDTs*. This is the first paper to study the inherent lower bounds of replicated data types. The contribution includes derivation of lower bounds for several data types, improvement of some implementations, and proved optimality of others (see Section 5.3.10). This result was presented at POPL 2014 [25].

5.2. Distributed algorithms for dynamic networks

Participants: Luciana Bezerra Arantes [correspondent], Rudyar Cortes, Raluca Diaconu, Jonathan Lejeune, Olivier Marin, Sébastien Monnet, Franck Petit [correspondent], Karine Pires, Pierre Sens, Véronique Simon, Julien Sopena.

Nowadays, distributed systems are more and more heterogeneous and versatile. Computing units can join, leave or move inside a global infrastructure. These features require the implementation of dynamic systems, that is to say they can cope autonomously with changes in their structure in terms of physical facilities and software. It therefore becomes necessary to define, develop, and validate distributed algorithms able to managed such dynamic and large scale systems, for instance mobile *ad hoc* networks, (mobile) sensor networks, P2P systems, Cloud environments, robot networks, to quote only a few.

Efficiency in such environments requires specialised protocols, providing features such as fault or heterogeneity tolerance, scalability, quality of service, and self-*. Our approach covers the whole spectrum from theory to experimentation. We design algorithms, prove them correct, implement them, and evaluate them in simulation, using OMNeT++ or PeerSim, and on large-scale real platforms such as Grid'5000. The theory ensures that our solutions are correct and whenever possible optimal; experimental evidence is necessary to show that they are relevant and practical.

Within this thread, we have considered a number of specific applications, including massively multi-player on-line games (MMOGs) and peer certification.

We have obtained results both on fundamental aspects of distributed algorithms and on specific emerging large-scale applications.

We study various key topics of distributed algorithms: mutual exclusion, failure detection, data dissemination and data finding in large scale systems, self-stabilization and self-* services.

5.2.1. Self-Stabilization.

We have also approached fault tolerance through self-stabilization. Self-stabilization is a versatile technique to design distributed algorithms that withstand transient faults.

In [43], we proposed a silent self-stabilizing leader election algorithm (SSLE, for short) for bidirectional connected identified networks of arbitrary topology. Starting from any arbitrary configuration, SSLE converges to a terminal configuration, where all processes know the ID of the leader, this latter being the process of minimum ID. Moreover, as in most of the solutions from the literature, a distributed spanning tree rooted at the leader is defined in the terminal configuration. This algorithm is written in the locally shared memory model. It assumes the distributed unfair daemon, the most general scheduling hypothesis of the model. Our algorithm requires no global knowledge on the network (such as an upper bound on the diameter or the number of processes, for example). We showed that its stabilization time is in $\Theta(n^3)$ steps in the worst case, where n is the number of processes. Its memory requirement is asymptotically optimal, *i.e.*, $\Theta(\log n)$ bits per processes. Its round complexity is of the same order of magnitude — *i.e.*, $\Theta(n)$ rounds — as the best existing algorithm designed with similar settings. To the best of our knowledge, this was the first self-stabilizing leader election algorithm for arbitrary identified networks that is proven to achieve a stabilization time polynomial in steps. By contrast, we show that the previous best existing algorithm designed with similar settings stabilizes in a non polynomial number of steps in the worst case.

We have also implemented SSLE in a high-level simulator to empirically evaluate its average performances. Experimental results tend to show that its worst case in terms of rounds ($\Theta(3n + D)$ rounds) is rare.

5.2.2. Dynamic Distributed Systems

The first key challenge in understanding highly dynamic networks consists in developing appropriate models that are as close as possible to the phenomena that one wishes to capture. This requires the use of a formalism sufficiently expressive to formulate complex temporal properties. Recently, a vast collection of concepts, formalisms, and models has been unified in a framework called Time-Varying Graphs (TVG) [0], which are represented as time-ordered sequences of graphs defined over a fixed set of nodes. A hierarchy of classes over TVG has been described, mainly depending on properties related to connectivity and recurrence of dynamic. Such an hierarchy is an interesting tool for study computability issues. As an example, if one is able to prove an impossibility result in a class of the hierarchy with strong properties, then this impossibility result also holds in any class of the hierarchy with (strictly) weaker properties. In this context, we provide a generic framework to prove impossibility results in this model [45]. This framework helps to formally prove classical arguments about convergence of sequence of time-varying graphs used to build counter-examples. We apply this generic framework to the study of covering problems (such as minimal dominating set and maximal matching) in the context of time-varying graphs. We obtain a characterization of the weakest topology assumption that makes these problems computable. We also propose a general time complexity measure since time-varying graph model lacks so far of such a definition.

5.2.3. Swarm of Mobile Robots

Swarm of autonomous mobile sensor devices (or, robots) recently emerged as an attractive issue in the study of dynamic distributed systems permits to assess the intrinsic difficulties of many fundamentals tasks, such as exploring or gathering in a discrete space. We consider autonomous robots that are endowed with visibility sensors (but that are otherwise unable to communicate) and motion actuators. Those robots must collaborate to solve a collective task, namely *exclusive perpetual exploration*, despite being limited with respect to input from the environment, asymmetry, memory, etc. The area to be explored is modeled as a graph and the exclusive perpetual exploration task requires every possible vertex to be visited infinitely often by every robot, with the additional constraint that no two robots may be present at the same node at the same time or may concurrently traverse the same edge of the graph.

In [28], we presented and implemented a generic method for obtaining all possible protocols for a swarm of mobile robots operating in a particular discrete space, namely an anonymous rings. Our method permits to discover new protocols that solve the problem, and to assess specific optimization criteria (such as individual coverage, visits frequency, etc.) that are met by those protocols. To our best knowledge, this was the first attempt to mechanize the discovery and fine-grained property testing of distributed mobile robot protocols.

[0] A. Casteigts, P. Flocchini, W. Quattrociocchi, and N. Santoro, Time-varying graphs and dynamic networks, International Journal of Parallel, Emergent and Distributed Systems 27(5):387-408, 2012

5.3. Management of distributed data

Participants: Pierpaolo Cincilla, Raluca Diaconu, Jonathan Lejeune, Mesaac Makpangou, Olivier Marin, Sébastien Monnet, Karine Pires, Dastagiri Reddy Malikireddy, Masoud Saeida Ardekani, Pierre Sens, Marc Shapiro, Véronique Simon, Julien Sopena, Vinh Tao Thanh, Serdar Tasiran, Marek Zawirski.

Storing and sharing information is one of the major reasons for the use of large-scale distributed computer systems. Replicating data at multiple locations ensures that the information persists despite the occurrence of faults, and improves application performance by bringing data close to its point of use, enabling parallel reads, and balancing load. This raises numerous issues:

- Where to store or replicate the data, in order to ensure that it is available quickly and remains persistent despite failures and disconnections.
- How many copies, located where, are needed to face dynamically-changing demand (load) and offer (elasticity).
- How to parallelize writes and hence how to ensure consistency between replicas.
- Tradeoffs between synchronised, consistent but slow updates, and fast but weakly-consistent ones.
- When and how to move data to computation, or computation to data, in order to improve response time while minimizing storage or energy usage.
- How to apply our approaches towards addressing the above issues onto a challenging use case: achieving true scalability for online games.

5.3.1. Long term durability

To tolerate failures, distributed storage systems replicate data. However, despite the replication, pieces of data may be lost (i.e. all the copies are lost). We have previously proposed a mechanism, RelaxDHT, to make distributed hash tables (DHT) resilient to high churn rates.

We have observed that a given system with a given replication mechanism can store a certain amount of data above which the loss rate would be greater than an "acceptable"/fixed threshold. This amount of data can be used as a metric to compare replication strategies. We have studied the impact of the data distribution layout upon the loss rate. The way the replication mechanism distribute the data copies among the nodes has a great impact. If node contents are very correlated, the number of available sources to heal a failure is low. On the opposite, if the data copies are shuffled/scattered among the nodes, many source nodes may be available to heal the system, and thus, the system losses less pieces of data. In order to study data durability on a long term, we have designed a model, and implemented a discrete event based simulator that can simulate a 100 node system over years within several hours. Our model, SPLAD [49] (for scattering and placing data replicas to enhance long-term durability), allows us to vary the data scattering degree by tuning a selection range width. We are also studying the impact of the policy used while choosing a storing node within the selection range (e.g., randomly, the least loaded, or smarter policies like the power of two choices). This policy has an important impact on both the storage load distribution among nodes and the number of lost pieces of data.

5.3.2. Achieving scalability for online games

Massively Multiplayer Online Games (MMOGs) such as *World of Warcraft* constitute a great use case for the management of distributed data on a large scale. Commercial support systems for MMOGs rely almost exclusively on traditional client/server architectures that are centralized. These architectures do not scale properly, both in terms of the number of players and of the number objects used to model virtual universes that grow ever more complex. Most MMOGs avoid this problem by limiting the scale of the universe: the virtual environment is partitioned into several parallel and totally disconnected worlds, such as the *Realms* in *World of Warcraft*. Each partition, handled in a centralized way, limits the number of players it can host; avatars created on different partitions will never meet in the game.

From a systems point of view, achieving true scalability raises many challenging issues for MMOGs. For instance the system must be very reactive: if the update latency on a player node is too high, the game becomes unplayable. Since these games are meant to operate on a large scale, they induce a trade-off between availability and consistency of data. The consistency aspect is critical because MMOGs incur a high degree of cheating.

Designing and implementing a scalable service for Multiplayer Online Games requires an extensive knowledge of the habits, behaviors and expectations of the players. The first part of our work on MMOGs aimed at gathering and analyzing traces of real games offers to gain insight on these matters. We collected public data from a *League of Legends* server (information over more than 56 million game sessions): the resulting database is freely available online, and an ensuing publication [34] details the analysis and conclusions we draw from this data regarding the expected requirements for a scalable MMOG service.

We steered a second part of our work on MMOGs in 2014 towards designing a peer to peer refereeing system that remains highly efficient, even on a large scale, both in terms of performance and in terms of cheat prevention. Simulations show that such a system scales easily to more than 30,000 nodes while leaving less than 0.013% occurrences of cheating undetected on a mean total of 24,819,649 refereeing queries. This work got published in the Multimedia Systems Journal [21].

Finally, we also worked on the design of a scalable architecture for online games. The goal is to balance the load among nodes to allow the simulation of a whole, contiguous, virtual space.

5.3.3. Management of dynamic big data

Managing and processing Dynamic Big Data, where multiple sources produce new data continuously, is very complex. Static cluster- or grid-based solutions are prone to induce bottleneck problems, and are therefore ill-suited in this context. Our objective in this domain is to design and implement a Reliable Large Scale Distributed Framework for the Management and Processing of Dynamic Big Data. In 2014, we focused our research on data placement and on gathering traces from target applications in order to assess our future solutions.

With respect to placement, we worked on a scheme to store and access massive streams of data efficiently. We designed a solution that extends distributed prefix tree indexing structures for this purpose. Our new maintenance protocol anticipates every data insertion on provisional child nodes and thus significantly reduces overhead and improves query response time. This work has lead to the publication of an Inria research report (RR- 8637) [46].

With respect to application traces, we targeted sport tracker applications. Designing and implementing a big data service for sport tracker applications requires an extensive knowledge of both data distribution and input load. Gathering and analysing traces from a real world sports tracker service provides insight on these matters, but such services are very protective of their data due to competition as well as privacy issues. We avoided these issues by gathering public data from a popular sports tracker server called EndoMondo. The resulting database is freely available online, and allowed an in-depth analysis from a dynamic big data perspective. This study has lead to the publication of an Inria research report (RR- 8636) [47] .

5.3.4. Adaptative replication

Different pieces of data have different popularity: some data are stored but never accessed while other pieces are very “hot” and are requested concurrently by many clients. This implies that different pieces of data with different popularity should have a different number of copies to efficiently serve the requests without wasting resources. Furthermore, for a given piece of data, the popularity may vary drastically among time. It is thus important that the replication mechanism dynamically adapt the number of replicas to the demand. In the context of the ODISEA2 FUI project, we have studied the popularity distribution and evolution of live video streams [31], [36].

5.3.5. Keyword-based Indexing and Search Substract for Structured P2P Information System

Number of large scale information systems rely on a DHT-based storage infrastructure. To help users to find suitable information, one attractive solution is to maintain an index that maps keywords to suitable data. Maintaining and exploiting an index distributed towards a DHT is confronted to the performance issue. Mainly, the computation of the intersection of postings related to provided keywords could generate too large traffic over the network; also one is confronted to some unbalanced on peers' load due to the fact that certain world are too popular!

In 2014, we propose *FreeCore*, a DHT-based distributed indexing substract that can be used to build efficient keyword-based search facilities for large scale information systems. A *FreeCore* index, considers keyword sets, then summarizes each set with a Bloom Filter. To limit the probability of false positive, we anticipate that one will use large size filters together enough hash functions. Thanks to this representation, we transform the searching problem, to the one of bitmaps matching as each query is also coded by a Bloom Filter. To distribute resulting summaries towards peers, *FreeCore* considers each summary as a sequence of binary keywords. Each binary keyword is assigned a peer and all summaries containing this binary keyword are stored at its assigned peer. Finally, to reduce the traffic overhead as well as the the size of local indices, *FreeCore* fragments each filter such as to factorize sequence of bits that occur more than once. In [40], we report the performances of the initial implementation of *FreeCore*. Thought a number of improvements were not included within this initial evaluation, *FreeCore* offers better performances than existing state of the art. Current work focusses on developping applications that exploit *FreeCore*.

5.3.6. Large-Scale File Systems

Storage architectures for large enterprises are evolving towards a hybrid cloud model, mixing private storage (pure SSD solutions, virtualization-on-premise) with cloud-based service provider infrastructures. Users will be able to both share data through the common cloud space, and to retain replicas in local storage. In this context we need to design data structures suitable for storage, access, update and consistency of massive amounts of data at the object, block or file system level.

Current designs consider only data structures (e.g., trees or B+-Trees) that are strongly consistent and partition-tolerant (CP). However, this means that they are not available when there is a network problem, and that replicating a CP index across sites is painful. The traditional approaches include locking, journaling and replaying of logs, snapshots and Merkle trees. All of these are difficult to scale using generic approaches, although it is possible to scale them in some specific instances. For instance, synchronization in a single direction (the Active/Passive model) is relatively simple but very limited. A multi-master (Active/Active) model, where updates are allowed at multiple replicas and synchronization occurs in both directions, is difficult to achieve with the above techniques.

Our previous work has shown that many storage indexing operations commute; this enables a the highly-scalable CRDT approach. For those that do not, the explicit consistency approach (Section 5.3.11) appears promising.

This work is part of a CIFRE agreeement with Scality (see Section 6.2.1).

5.3.7. Strong consistency

When data is updated somewhere on the network, it may become inconsistent with data elsewhere, especially in the presence of concurrent updates, network failures, and hardware or software crashes. A primitive such as consensus (or equivalently, total-order broadcast) synchronises all the network nodes, ensuring that they all observe the same updates in the same order, thus ensuring strong consistency. However the latency of consensus is very large in wide-area networks, directly impacting the response time of every update. Our contributions consist mainly of leveraging application-specific knowledge to decrease the amount of synchronisation.

When a database is very large, it pays off to replicate only a subset at any given node; this is known as partial replication. This allows non-overlapping transactions to proceed in parallel at different locations and decreases the overall network traffic. However, this makes it much harder to maintain consistency. We designed and implemented two *genuine* consensus protocols for partial replication, i.e., ones in which only relevant replicas participate in the commit of a transaction.

Another research direction leverages isolation levels, particularly Snapshot Isolation (SI), in order to parallelize non-conflicting transactions on databases. We prove a novel impossibility result: under standard assumptions (data store accesses are not known in advance, and transactions may access arbitrary objects in the data store), it is impossible to have both SI and GPR. Our impossibility result is based on a novel decomposition of SI which proves that, like serializability, SI is expressible on plain histories.

We designed an efficient protocol that maintains side-steps this impossibility but maintains the most important features of SI:

1. (Genuine Partial Replication) only replicas updated by a transaction T make steps to execute T;
2. (Wait-Free Queries) a read-only transaction never waits for concurrent transactions and always commits;
3. (Minimal Commit Synchronization) two transactions synchronize with each other only if their writes conflict.

The protocol also ensures Forward Freshness, i.e., that a transaction may read object versions committed after it started.

Non-Monotonic Snapshot Isolation (NMSI) is the first strong consistency criterion to allow implementations with all four properties. We also present a practical implementation of NMSI called Jessy, which we compare experimentally against a number of well-known criteria. Our measurements show that the latency and throughput of NMSI are comparable to the weakest criterion, read-committed, and between two to fourteen times faster than well-known strong consistencies.

An interesting side-effect of this research is an apples-to-apples comparison of many strong-consistency protocols. This work was published at LADIS 2014 [41] and at Middleware 2014 [33].

This research is supported in part by ConcoRDanT ANR project (Section 7.1.7) and by the FP7 grant SyncFree (Section 7.2.1.1).

5.3.8. *Distributed Transaction Scheduling*

Parallel transactions in distributed DBs incur high overhead for concurrency control and aborts. Our Gargamel system proposes an alternative approach by pre-serializing possibly conflicting transactions, and parallelizing non-conflicting update transactions to different replicas. This system provides strong transactional guarantees. In effect, Gargamel partitions the database dynamically according to the update workload. Each database replica runs sequentially, at full bandwidth; mutual synchronisation between replicas remains minimal. Both our simulations and the experimental results obtained with our prototype show that Gargamel improves both response time and load by an order of magnitude when contention is high (highly loaded system with bounded resources), and that otherwise slow-down is negligible.

We have studied Gargamel's behavior while running over multiple geographically distant sites. One instance of Gargamel runs on each site, synchronizations among the different sites occur off the critical path [39]. Our experiments with the Amazon platform show that or solution can be used to support failures of whole sites.

5.3.9. *Eventual consistency*

Eventual Consistency (EC) aims to minimize synchronisation, by weakening the consistency model. The idea is to allow updates at different nodes to proceed without any synchronisation, and to propagate the updates asynchronously, in the hope that replicas converge once all nodes have received all updates. EC was invented for mobile/disconnected computing, where communication is impossible (or prohibitively costly). EC also appears very appealing in large-scale computing environments such as P2P and cloud computing. However, its apparent simplicity is deceptive; in particular, the general EC model exposes tentative values, conflict resolution, and rollback to applications and users. Our research aims to better understand EC and to make it more accessible to developers.

We propose a new model, called *Strong Eventual Consistency* (SEC), which adds the guarantee that every update is durable and the application never observes a roll-back. SEC is ensured if all concurrent updates have a deterministic outcome. As a realization of SEC, we have also proposed the concept of a Conflict-free Replicated Data Type (CRDT). CRDTs represent a sweet spot in consistency design: they support concurrent updates, they ensure availability and fault tolerance, and they are scalable; yet they provide simple and understandable consistency guarantees.

This new model is suited to large-scale systems, such as P2P or cloud computing. For instance, we propose a "sequence" CRDT type called Treedoc that supports concurrent text editing at a large scale, e.g., for a wikipedia-style concurrent editing application. We designed a number of CRDTs such as counters (supporting concurrent increments and decrements), sets (adding and removing elements), graphs (adding and removing vertices and edges), and maps (adding, removing, and setting key-value pairs).

CRDTs are the main topic of the ConcoRDanT ANR project (Section 7.1.7) and the FP7 grant SyncFree (Section 7.2.1.1). After developing the SwiftCloud extreme-scale CRDT platform (see Section 4.3), we are currently developing a flexible cloud database called Antidote (see Section 4.4).

5.3.10. Lower bounds and optimality of CRDTs

CRDTs raise challenging research issues: What is the power of CRDTs? Are the sufficient conditions necessary? How to engineer interesting data types to be CRDTs? How to garbage collect obsolete state without synchronisation, and without violating the monotonic semi-lattice requirement? What are the upper and lower bounds of CRDTs?

We co-authored an innovative approach to these questions, published at Principles of Programming Languages (POPL) 2014 [25]. Geographically distributed systems often rely on replicated eventually consistent data stores to achieve availability and performance. To resolve conflicting updates at different replicas, researchers and practitioners have proposed specialized consistency protocols, called replicated data types, that implement objects such as registers, counters, sets or lists. Reasoning about replicated data types has however not been on par with comparable work on abstract data types and concurrent data types, lacking specifications, correctness proofs, and optimality results. To fill in this gap, we propose a framework for specifying replicated data types using relations over events and verifying their implementations using replication-aware simulations. We apply it to seven existing implementations of 4 data types with nontrivial conflictresolution strategies and optimizations (last-writer-wins register, counter, multi-value register and observed-remove set). We also present a novel technique for obtaining lower bounds on the worst-case space overhead of data type implementations and use it to prove optimality of four implementations. Finally, we show how to specify consistency of replicated stores with multiple objects axiomatically, in analogy to prior work on weak memory models. Overall, our work provides foundational reasoning tools to support research on replicated eventually consistent stores.

5.3.11. Explicit Consistency: Strengthening Eventual Consistency to support application invariants

The designers of the replication protocols for geo-replicated storage systems have to choose between either supporting low latency, eventually consistent operations, or supporting strong consistency for ensuring application correctness. We propose an alternative consistency model, *explicit consistency*, that strengthens eventual consistency with a guarantee to preserve specific invariants defined by the applications. Given these application-specific invariants, a system that supports explicit consistency must identify which operations are unsafe under concurrent execution, and help programmers to select either violation-avoidance or invariant-repair techniques. We show how to achieve the former while allowing most of operations to complete locally, by relying on a reservation system that moves replica coordination off the critical path of operation execution. The latter, in turn, allow operations to execute without restriction, and restore invariants by applying a repair operation to the database state. We designed and evaluated Indigo, a middleware that provides Explicit

Consistency on top of a causally-consistent data store. Indigo guarantees strong application invariants while providing latency similar to an eventually consistent system.

This work was presented at W-PSDS 2014 [24] and LADIS 2014 [38]. It was selected for presentation at EuroSys 2015 [23]. This research is supported in part by the FP7 grant SyncFree (Section 7.2.1.1).

5.4. Memory management for big data

Participants: Antoine Blin, Lokesh Gidra, Sébastien Monnet, Marc Shapiro, Julien Sopena [correspondent], Gaël Thomas.

5.4.1. Garbage collection for big data on large-memory NUMA machines

On contemporary cache-coherent Non-Uniform Memory Access (ccNUMA) architectures, applications with a large memory footprint suffer from the cost of the garbage collector (GC), because, as the GC scans the reference graph, it makes many remote memory accesses, saturating the interconnect between memory nodes. We address this problem with NumaGiC, a GC with a mostly-distributed design. In order to maximise memory access locality during collection, a GC thread avoids accessing a different memory node, instead notifying a remote GC thread with a message; nonetheless, NumaGiC avoids the drawbacks of a pure distributed design, which tends to decrease parallelism. We compared NumaGiC with Parallel Scavenge and NAPS on two different ccNUMA architectures running on the Hotspot Java Virtual Machine of OpenJDK 7. On Spark and Neo4j, two industry-strength analytics applications, with heap sizes ranging from 160 GB to 350 GB, and on SPECjbb2013 and SPECjbb2005, NumaGiC improves overall performance by up to 45% over NAPS (up to 94% over Parallel Scavenge), and increases the performance of the collector itself by up to 3.6× over NAPS (up to 5.4× over Parallel Scavenge).

This research is accepted for presentation at the ASPLOS 2015 conference [29].

5.4.2. File cache pooling

Some applications, like online sales servers, intensively use disk I/Os. Their performance is tightly coupled with I/Os efficiency. To speed up I/Os, operating systems use free memory to offer caching mechanisms. Several I/O intensive applications may require a large cache to perform well. However, nowadays resources are virtualized. In clouds, for instance, virtual machines (VMs) offer both isolation and flexibility.This is the foundation of cloud elasticity, but it induces fragmentation of the physical resources, including memory. This fragmentation reduces the amount of available memory a VM can use for caching I/Os. We propose Puma [35] (for Pooling Unused Memory in Virtual Machines) which allows I/O intensive applications running on top of VMs to benefit of large caches.

This is realized by providing a remote caching mechanism that provides the ability for any VM to extend its cache using the memory of other VMs located either in the same or in a different host. Puma is a kernel level remote caching mechanism that is: (*i*) block device, file system and hypervisor agnostic; and (*ii*) efficient both locally and remotely. It can increase applications performance up to 3 times without impacting potential activity peaks.

REGULARITY Project-Team

6. New Results

6.1. Highlights of the Year

The article "Christiane's Hair" by Jacques Lévy-Véhel and Franklin Mendivil has received the Paul R. Halmos - Lester R. Ford award of the Mathematical Association of America.

6.2. Modelling the exchange of cultural goods on the Internet

Participant: Jacques Lévy Véhel.

In collaboration with Pierre Emmanuel Lévy Véhel and Victor Lévy Véhel.

Illegal sharing of cultural goods on the Internet has become a massive reality in today's connected society. Numerous studies have been performed to try and evaluate the impact of these practices on the industry of cultural goods, and how much harm, if any, they have entailed. The effect of legal and technical responses to limit pirating has also been investigated, showing in general inconclusive effect. Instead of penalizing illegal actors - providers and/or consumers -, a totally different approach has been proposed recently by the french government agency Hadopi. The idea is to offer the possibility to sites that illegally share cultural goods to become legal in exchange of a retribution proportional to their activity. In the frame of a contract with the Hadopi, we have built a model that studies the economic feasibility of such a scheme under various assumptions on the behaviour of the different actors involved. Our main finding is that, supposing that more popular goods are more prone to pirating, a retribution of the order of the increase in benefit per user gained by legalized sites does indeed lead to a win-win situation for both producers/sellers of cultural goods and willing-to-be-legalized sites. This will be the case under two conditions: the proportion of pirates is large enough (which seems largely true) and the increase in the amount of money that forums will make from advertisement when becoming legal is sufficient [43].

An extension of our work is under way, that will consider further actors and refined modelling of the way illegal sharing takes place. Calibration issues will also be investigated more closely.

6.3. Financial risk analysis

Participant: Jacques Lévy Véhel.

Financial regulations have fundamentally changed since the Basel II Accords. Among other evolutions, Basel II and III explicitly impose that computations of capital requirements be model-based. This paradigm shift in risk management has been the source of strong debates among both practitioners and academics, who question whether such model-based regulations are indeed more efficient.

A common feeling in the industry is that regulations will sometimes give a false impression of security: risk manager tend to think that a financial company that would fulfil all the criteria of, say, the Basel III Accords on capital adequacy, is not necessarily on the safe side. This is so mainly because many risks, and most significantly systemic or system-wide risks, are not properly modelled, and also because it is easy to manipulate to some extent various risk measures, such as VaR.

In parallel, a fast growing body of academic research provides various arguments explaining why current regulations are not well fitted to address risk management in an adequate way, and may even, in certain cases, worsen the situation.

We use the term *regulation risk* to describe the fact that, in some situations, prudential rules are themselves the source of a systemic risk. We have shown how a combination of model risk and regulation risk leads to an effect which is exactly the opposite of what the regulator tries to enforce. More precisely, we explain how wrongly assuming a Gaussian dynamics (or, more generally, a left-light-tailed one) when the "true" one is pure jump (or, more generally, left-heavy-tailed), and imposing as a constraint *minimizing* VaR at constant volume results in effect in movements that will *maximize* VaR. This effect is related to the fact that regulations fail to consider that risk is endogenous. In a nutshell, the idea is simply that, by treating jumps in the evolution of prices as exceptional events and essentially ignoring them in model-based VaR computations, one misses an essential dimension of risk, and acts in a way that will in effect favour sudden large movements in the markets and ultimately increase VaR. Our simple setting predicts that VaR constraints result in an *increased* intensity of jumps and a *decrease* in volatility - a fact confirmed experimentally on certain datasets. This is a mathematical translation of the common feeling of practitioners that regulations give a false impression of security characterized by low volatility but increased risk of sudden large movements.

6.4. Functional central limit theorem for multistable Lévy motions

Participants: Xiequan Fan, Jacques Lévy Véhel.

We prove a functional central limit theorem (FCLT) for the independent-increments multistable Lévy motions (MsLM) $L_I(t), t \in [0,1]$, as well as of integrals with respect to these processes, using weighted sums of independent random variables. In particular, we prove that multistable Lévy motions are stochastic Hölder continuous and strongly localisable.

Theorem 0.1 Let $(\alpha_n(u))_n, \alpha(u), u \in [0,1]$, be a class of càdlàg functions ranging in $[a,b] \subset (0,2]$ such that the sequence $(\alpha)_n$ tends to α in the uniform metric. Let $(X(k,n))_{n\in\mathbb{N},\ k=1,\ldots,2^n}$ be a family of independent and symmetric $\alpha_n(\frac{k}{2^n})-$stable random variables with unit scale parameter, i.e., $X(k,n) \sim S_{\alpha_n(\frac{k}{2^n})}(1,0,0)$. Then the sequence of processes

$$L_I^{(n)}(u) = \sum_{k=1}^{\lfloor 2^n u \rfloor} \left(\frac{1}{2^n}\right)^{1/\alpha_n(\frac{k}{2^n})} X(k,n), \qquad u \in [0,1], \tag{17}$$

tends in distribution to $L_I(u)$ in $(D[0,1], d_S)$, where $\lfloor x \rfloor$ is the largest integer smaller than or equal to x. In particular, if α satisfies

$$(\alpha(x) - \alpha(x+t)) \ln t \to 0 \tag{18}$$

uniformly for all x as $t \searrow 0$, then $L_I(u)$ is localisable at all times.

We have defined integrals of MsLM, and given criteria for convergence,independence, stochastic Hölder continuity and strong localisability of such integrals.

6.5. Deviation inequalities for martingales with applications

Participant: Xiequan Fan.

In the papers [36], [37] we study some general exponential inequalities for supermartingales. The inequalities improve or generalize many exponential inequalities of Bennett (1962), Freedman (1975), van de Geer (1995), de la Peña (1999) and Pinelis (2006). Moreover, our concentration inequalities also improve some known inequalities for sums of independent random variables. Applications associated with linear regressions, autoregressive processes and branching processes are provided. In particular, an interesting application of de la Peña's inequality to self-normalized deviations is also provided.

We also considered an $\mathcal{X}$-valued Markov chain $X_1, X_2, ..., X_n$ belonging to a class of iterated random functions, which is “one-step contracting" with respect to some distance d on $\mathcal{X}$. If f is any separately Lipschitz function with respect to d, we use a well known decomposition of $S_n = f(X_1, ..., X_n) - \mathbb{E}[f(X_1, ..., X_n)]$ into a sum of martingale differences d_k with respect to the natural filtration $\mathcal{F}_k$. We show that each difference d_k is bounded by a random variable η_k independent of $\mathcal{F}_{k-1}$. Using this very strong property, we obtain a large variety of deviation inequalities for S_n, which are governed by the distribution of the η_k's. Finally, we give an application of these inequalities to the Wasserstein distance between the empirical measure and the invariant distribution of the chain.

6.6. Self-stabilizing Lévy motions

Participants: Xiequan Fan, Jacques Lévy Véhel.

Self-stabilizing processes have the property that the “local intensities of jumps” varies with amplitude. They are good models for, e.g., financial and temperature records.

The main aim of our work is to establish the existence of such processes and to give a simple construction. Formally, one says that a stochastic process $S(t), t \in [0,1]$, is a self-stabilizing process if, for almost surely all $t \in [0,1)$, S is localisable at t with tangent process S'_t an $g(S(t))-$stable process, with respect to the conditional probability measure $\mathbb{P}_{S(t)}$. In other words,

$$\lim_{r \searrow 0} \frac{S(t+ru) - S(t)}{r^{1/g(S(t))}} = S'_t(u), \tag{19}$$

where convergence is in finite dimensional distributions with respect to $\mathbb{P}_{S(t)}$. Heuristically, if $S'_t(u) = L_{g(S(t))}(u)$, equality (8) implies that

$$S(t+ru) - S(t) \approx r^{1/g(S(t))} L_{g(S(t))}(u) = (ru)^{1/g(S(t))} L_{g(S(t))}(1),$$

when r is small. Thus it is natural to define $S(t) = \lim_{n\to\infty} S_n(\frac{\lfloor nt \rfloor}{n})$, where

$$S_n\left(\frac{k+1}{n}\right) - S_n\left(\frac{k}{n}\right) = n^{-1/g(S_n(k/n))} L_{g(S_n(k/n))}(1).$$

This inspiration allows us to build Markov processes that converge to a self-stabilizing process. Note that, when $\alpha(x) \equiv 2$, this is simply Donsker's construction. The main difficult is to prove the weak convergence of S_n. To this aim, we make use of a generalization of the Arzelà-Ascoli theorem.

Definition 0.1 We call the sequence $(f_n(\theta))_{n\geq 1}$ is sub-equicontinuous on $I \subset \mathbb{R}^d$, if for any $\varepsilon > 0$, there exist $\delta > 0$ and a sequence of nonnegative numbers $(\varepsilon_n)_{n\geq 1}$,$\varepsilon_n \to 0$ as $n \to \infty$, such that, for all functions f_n in the sequence,

$$|f_n(\theta_1) - f_n(\theta_2)| \ \leq \ \varepsilon + \varepsilon_n, \qquad \theta_1, \theta_2 \in I, \tag{20}$$

whenever $||\theta_1 - \theta_2|| < \delta$ (if $\varepsilon_n = 0$ for all n, then $(f_n(\theta))_{n\geq 1}$ is just equicontinuous).

The slightly generalized version of the Arzelà-Ascoli theorem reads:

Lemma 0.1 Assume that $(f_n)_{n\geq 1}$ be a sequence of real-valued continuous functions defined on a closed and bounded set $\Pi_{i=1}^d [a_i, b_i] \subset \mathbb{R}^d$. If this sequence is uniformly bounded and sub-equicontinuous, then there exists a subsequence $(f_{n_k})_{k\geq 1}$ that converges uniformly.

The following theorem states that self-stabilizing processes do exist.

Theorem 0.2 Let g be a Hölder function defined on $\mathbb{R}$ and ranging in $[a,b] \subset (0,2]$. There exists a self-stabilizing process $S(t), t \in [0,\, 1]$, that it is tangent at all u to a $g(S(u))-$stable Lévy process under the conditional expectation with respect to $S(u)$. Moreover, the process $S(t), t \in [0,\, 1]$, satisfies, for all $(\theta_j, t_j) \in \mathbb{R} \times [0,1]$, $j = 1, 2, ..., d$,

$$\mathbb{E}_{S(t_1)} \left[\exp \left\{ i \sum_{j=2}^{d} \theta_j \left(S(t_j) - S(t_1) \right) + \int | \sum_{j=2}^{d} \theta_j \mathbf{1}_{[t_1, t_j]}(z) |^{g(S(z))} dz \right\} \right] = 1. \tag{21}$$

We are currently studying the main properties of self-stabilizing processes.

REO Project-Team

6. New Results

6.1. Highlights of the Year

- Jimmy Mullaert was awarded the best poster prize at the conference Canum 2014.
- Jessica Oakes was awarded a University of California Presidential Postdoctoral Fellowship.
- Jessica Oakes won a young investigator award at the “4th International Conference on Engineering Frontiers in Pediatric and Congenital Heart Disease”.

6.2. Mathematical and numerical analysis of fluid-structure interaction problems

Participants: Benoit Fabrèges, Miguel Ángel Fernández Varela, Mikel Landajuela Larma, Jimmy Mullaert, Marina Vidrascu.

- In [54] we introduce two new classes of numerical methods for the solution of incompressible fluid/thin-walled structure interaction problems with unfitted meshes. The semi-implicit or explicit nature of the splitting in time is dictated by the order in which the spatial and time discretizations are performed. Stability and optimal accuracy are achieved without restrictive CFL conditions or correction iterations. Results presented by M. Landajuela at the 11th World Congress on Computational Mechanics (WCCM XI), July 20-25, 2014, Barcelona (Spain).
- In [47] we introduce a class of fully decoupled time-marching schemes (velocity-pressure-displacement splitting) for the coupling of an incompressible fluid with a thin-walled viscoelastic structure. The time splitting combines a projection method in the fluid with a specific Robin-Neumann treatment of the interface coupling. A priori energy estimates guaranteeing unconditional stability are established for some of the schemes. The accuracy and performance of the methods proposed is illustrated by a thorough numerical study.
- We have performed an a priori error analysis for the generalized Robin-Neumann explicit coupling schemes introduced in [30]. The analysis confirms the $\mathcal{O}(\tau^{2^{r-1}}/h^{\frac{1}{2}})$ error perturbation anticipated by the numerical evidence of [30]. Another fundamental result of this work is that the h-non-uniformity of the splitting error is not a consequence of the mass-lumping approximation (which simply dictates the explicit or semi-implicit nature of the coupling scheme). The analysis indicates that the genesis of the $\mathcal{O}(h^{-\frac{1}{2}})$ is the non-uniformity of discrete viscoelastic operator, which is a consequence of thick-walled character of the solid. These results have been reported in [48] and presented by M.A. Fernández at the 11th World Congress on Computational Mechanics (WCCM XI), July 20-25, 2014, Barcelona (Spain).
- We consider the extension of the Nitsche-XFEM method to fluid-structure interaction problems involving a thin-walled elastic structure (Lagrangian formalism) immersed in an incompressible fluid (Eulerian formalism). The fluid domain is discretized with an unstructured mesh not fitted to the solid mid- surface mesh. Weak and strong discontinuities across the interface are allowed for the velocity and pressure, respectively. The kinematic/kinetic fluid-solid coupling is enforced consistently using a variant of Nitsche’s method involving cut elements. Robustness with respect to arbitrary interface/element intersections is guaranteed through a ghost penalty stabilization. Different coupling schemes, either fully implicit or loosely coupled, are proposed. Several numerical examples, involving static and moving interfaces, illustrate the performance of the methods. A paper in collaboration with F. Alauzet (project-team Gamma3) is under preparation. Results presented by B. Fabrèges at the 11th World Congress on Computational Mechanics (WCCM XI), July 20-25, 2014, Barcelona (Spain).

6.3. Numerical methods for biological flows

Participants: Grégory Arbia, Benoit Fabrèges, Miguel Ángel Fernández Varela, Justine Fouchet-Incaux, Jean-Frédéric Gerbeau, Céline Grandmont, Sanjay Pant, Saverio Smaldone, Marc Thiriet, Irène Vignon-Clementel.

- In [19] We consider the problem of estimating the stiffness of an artery wall using a data assimilation method applied to a 3D fluid-structure interaction (FSI) model. We briefly present the FSI model, the data assimilation procedure based on a reduced order Unscented Kalman filter, and the segmentation algorithm. We then present two examples of the procedure using real data. First, we estimate the stiffness distribution of a silicon rubber tube from image data. Second, we present the estimation of aortic wall stiffness from real clinical data.
- In [29], we propose a new approach to the loosely coupled time-marching of a fluid-fluid interaction problems involving the incompressible Navier-Stokes equations. The methods combine a specific explicit Robin-Robin treatment of the interface coupling with a weakly consistent interface pressure stabilization in time. A priori energy estimates guaranteeing stability of the splitting are obtained for a total pressure formulation of the coupled problem. The performance of the proposed schemes is illustrated on several numerical experiments related to simulation of aortic blood flow.
- In [55] we investigate the stability of numerical schemes that are classically used in the simulation of airflows and blood flows. The geometrical complexity of the networks in which air/blood flows leads to a classical decomposition of two areas: a truncated 3D geometry corresponding to the largest contribution of the domain, and a 0D part connected to the 3D part, modelling air/blood flows in smaller airways/vessels. The resulting Navier-Stokes system in the 3D truncated part may involve non-local boundary conditions, deriving from a mechanical model. For various 3D/0D coupled models, different discretization processes are presented and analyzed in terms of numerical stability, highlighting strong differences according to the regimes that are considered. In particular, two main stability issues are investigated: first the coupling between the 3D and the 0D part for which implicit or explicit strategies are studied and, second, the question of estimating the amount of kinetic energy entering the 3D domain because of the artificial boundaries. The second issue has been also the subject of a review [31].
- In [31] we deal with numerical simulations of incompressible Navier-Stokes equations in truncated domain. In this context, the formulation of these equations has to be selected carefully in order to guarantee that their associated artificial boundary conditions are relevant for the considered problem. In this paper, we review some of the formulations proposed in the literature, and their associated boundary conditions. Some numerical results linked to each formulation are also presented. We compare different schemes, giving successful computations as well as problematic ones, in order to better understand the difference between these schemes and their behaviours dealing with systems involving Neumann boundary conditions. We also review two stabilization methods which aim at suppressing the instabilities linked to these natural boundary conditions.
- In [40], we propose a framework for Windkessel parameter estimation in a 0D representation of the 3D fluid-flow domain. Parameters are estimated from uncertain measurements through a sequential approach, and the 0D representation is iteratively improved through 3D-CFD simulations. The application of generalized sensitivity functions to assess parameter correlation and to ascertain the measurement set needed to avoid identifiability problems is also presented through representative test cases. This method, which is capable of handling non-simultaneous measurements, is demonstrated and validated for a patient-specific case of aortic coarctation.
- In [17] we perform the first patient-specific pulmonary hemodynamics 3D-0D modeling before single ventricle stage 2 surgery. 0D parameters are automatically tuned to match flow and pressure clinical measurements that are not taken where 3D boundary conditions need to be specified. This work on six patients demonstrates how simulations can help to check the coherence of clinical data or provide insights to clinicians that are otherwise difficult to measure, such as in the presence of kinks.

- In [25] we study a case of post single ventricle stage 2 surgery with the three following aims: (i) to show how to build a patient-specific model describing the hemodynamics in the presence of collaterals, using patient-specific clinical data collected at different times; (ii) to use this model to perform virtual collateral occlusion for quantitative hemodynamics prediction; and (iii) to compare predicted hemodynamics with post-operative measurements.

6.4. Numerical methods for cardiac electrophysiology

Participants: Muriel Boulakia, Jean-Frédéric Gerbeau, Damiano Lombardi, Elisa Schenone.

- In [33], a reduced-order method based on Approximated Lax Pairs (ALP) is applied to the integration of electrophysiology models. These are often high- dimensional parametric equation systems, challenging from a model reduction stand- point. The method is tested on two and three dimensional test-cases, of increasing complexity. The solutions are compared to the ones obtained by a finite element. The reduced-order simulation of pseudo-electrocardiograms based on ALP is proposed in the last part.
- In [21], we address the question of the discretization of Stochastic Partial Differential Equations (SPDE) for excitable media. Working with SPDE driven by colored noise, we consider a numerical scheme based on finite differences in time (Euler-Maruyama) and finite elements in space. Motivated by biological considerations, we study numerically the emergence of reentrant patterns in excitable systems such as the Barkley or Mitchell-Schaeffer models.

6.5. Lung and respiration modeling

Participants: Laurent Boudin, Muriel Boulakia, Céline Grandmont, Jessica Oakes, Ayman Moussa, Irène Vignon-Clementel.

- In [20], we consider the non-reactive fully elastic Boltzmann equation for mixtures. We deduce that, under the standard diffusive scaling, its limit for vanishing Mach and Knudsen numbers is the Maxwell-Stefan model for a multicomponent gaseous mixture.
- In [49], we first deal with the modelling and the discretization of an aerosol evolving in the air, in the respiration framework, within a domain which can be fixed or moving. We also investigate basic numerical properties of the numerical code which was developped, and also focus on the influence of the aerosol on the airflow.
- In [38], the aim of the study was to determine susceptibility differences between healthy and emphysematous rats exposed to airborne particles. To do this, we performed animal exposure experimenters and measured particle deposition concentrations with Magnetic Resonance Imaging. We showed that overall deposition was significantly higher in the elastase-treated rats compared to the healthy ones, suggesting enhanced susceptibility to airborne particles in diseased lungs. Current work aims at integrating such experimental data into modeling [39] and compare numerical simulations with experiments. To extend particle modeling to expiration, a 1D particle transport model is under development [44].

  While it is known that the retention of fine particles is less in microgravity (uG) compared to normal gravity (1G) levels, it was unknown the spatial relationship of deposited particles. In [26], rats were exposed to 1 micron diameter particles on the NASA uG airplane and compared to rats exposed in 1G. We found that the ratio of deposited particles in the central airways compared to the peripheral ones, was significantly less in the uG than in 1G, indicating enhanced deposition in the periphery. This data suggests that toxicology effects of exposure to Moon dust may not be insignificant.
- In [51], we establish stability estimates for the unique continuation property of the nonstationary Stokes problem. These estimates hold without prescribing boundary conditions and are of logarithmic type. They are obtained thanks to Carleman estimates for parabolic and elliptic equations. Then, these estimates are applied to an inverse problem where we want to identify a Robin coefficient defined on some part of the boundary from measurements available on another part of the boundary.

6.6. Miscellaneous

Participants: Jean-Frédéric Gerbeau, Damiano Lombardi, Marina Vidrascu.

- in [32] a reduced-order model algorithm, called ALP, is proposed to solve nonlinear evolution partial differential equations. It is based on approximations of generalized Lax pairs. Contrary to other reduced-order methods, like Proper Orthogonal Decomposition, the basis on which the solution is searched for evolves in time according to a dynamics specific to the problem. It is therefore well-suited to solving problems with progressive front or wave propagation. Another difference with other reduced-order methods is that it is not based on an off-line / on-line strategy. Numerical examples are shown for the linear advection, KdV and FKPP equations, in one and two dimensions.
- in [41] we propose a direct method for computing modal coupling coefficients - due to geometrically nonlinear effects - for thin shells vibrating at large amplitude and discretized by a finite element (FE) procedure. These coupling coefficients arise when considering a discrete expansion of the unknown displacement onto the eigenmodes of the linear operator. The evolution problem is thus projected onto the eigenmodes basis and expressed as an assembly of oscillators with quadratic and cubic nonlinearities. The nonlinear coupling coefficients are directly derived from the finite element formulation, with specificities pertaining to the shell elements considered, namely, here elements of the "Mixed Interpolation of Tensorial Components" family (MITC). Therefore, the computation of coupling coefficients, combined with an adequate selection of the significant eigenmodes, allows the derivation of effective reduced-order models for computing - with a continuation procedure - the stable and unstable vibratory states of any vibrating shell, up to large amplitudes. The procedure is illustrated on a hyperbolic paraboloid panel. Bifurcation diagrams in free and forced vibrations are obtained. Comparisons with direct time simulations of the full FE model are given. Finally, the computed coefficients are used for a maximal reduction based on asymptotic nonlinear normal modes (NNMs), and we find that the most important part of the dynamics can be predicted with a single oscillator equation.
- in [53] we deal with the following data assimilation problem: construct an analytical approximation of a numerical constitutive law in three-dimensional nonlinear elasticity. More precisely we are concerned with a micro-macro model for rubber as the one proposed in [36]. Macroscopic quantities of interest such as the Piola-Kirchhoff stress tensor can be approximated for any value of the strain gradient by numerically solving a nonlinear PDE. This procedure is however computationally demanding. Hence, although conceptually satisfactory, this physically-based model is of no direct practical use. We aim to circumvent this difficulty by proposing a numerical strategy to reconstruct from in silico experiments an accurate analytical proxy for the micro-macro constitutive law.

REVES Project-Team

6. New Results

6.1. Highlights of the Year

Our work on sketch-based modeling for product designers (Sec. 6.4.4) has received significant attention. It appeared on the news page of University of British Columbia http://news.ubc.ca/2014/08/13/powerful-math-creates-3-d-shapes-from-simple-sketches/ and our video has been watched more than 7000 times on Youtube http://youtu.be/tbUljHJv4Rg. We filed a patent on this technology and we have contacts with several companies about a potential transfer.

Our poster on *C-LOD: Context-aware Material Level-of-Detail applied to Mobile Graphics* [] received the 3rd place in the ACM's Graduate Student Research Competition at SIGGRAPH 2014. This work is a collaboration with George Alex Koulieris and Katerina Mania from the Technical University of Crete and Douglas Cunningham from the Technical University of Cottbus.
BEST PAPER AWARD :
[] **Computer Graphics Forum**. G. A. KOULIERIS, G. DRETTAKIS, D. W. CUNNINGHAM, K. MANIA.

6.2. Plausible and Realistic Image Rendering

6.2.1. Multi-View Intrinsic Images for Outdoors Scenes with an Application to Relighting

Participants: Sylvain Duchêne, Clement Riant, Gaurav Chaurasia, Stefan Popov, Adrien Bousseau, George Drettakis.

We introduce a method to compute intrinsic images for a multi-view set of outdoor photos with cast shadows, taken under the same lighting. We use an automatic 3D reconstruction from these photos and the sun direction as input and decompose each image into reflectance and shading layers, despite the inaccuracies and missing data of the 3D model. Our approach is based on two key ideas. First, we progressively improve the accuracy of the parameters of our image formation model by performing iterative estimation and combining 3D lighting simulation with 2D image optimization methods. Second we use the image formation model to express reflectance as a function of discrete visibility values for shadow and light, which allows us to introduce a robust visibility classifier for pairs of points in a scene. This classifier is used for shadow labeling, allowing us to compute high quality reflectance and shading layers. We then create shadow-caster geometry that preserves shadow silhouettes. Combined with the intrinsic layers, this approach allows multi-view relighting with moving cast shadows. We present results on several multi-view datasets, and show how it is now possible to perform image-based rendering with changing illumination conditions.

This work is part of an industrial partnership with Autodesk and is under revision for ACM Transactions On Graphics.

6.2.2. Compiler and Tiling Strategies for IIR Filters

Participants: Gaurav Chaurasia, George Drettakis.

We present a compiler for parallelizing IIR or recursive filters. IIR filters are frequently used for $O(1)$ convolutions, but they cannot exploit GPUs because they are very hard to parallelize and also exhibit poor memory locality which hinders performance on both CPUs and GPUs. We present algorithmic tiling strategies for IIR filters which overcome these limitations. Tiled IIR filters are notoriously hard to implement and hence largely ignored by programmers and hardware vendors. We present a compiler front-end that supports intuitive functional specification and tiling of IIR filters. We demonstrate that different tiling strategies may be optimal on different platforms and filter parameters; our compiler can express the exhaustive set of alternatives in just 10-20 lines of code. This enables programmers to easily explore a large variety of trade-offs at different levels of granularity, thereby making it easier and more likely to discover the optimal implementation, while also producing intuitive and maintainable code. Our initial results show that our compiler is as terse as vendor provided libraries, but it allows exploiting the algorithmic advantages of tiling which cannot be provided by any precompiled library.

For example, our compiler can compute a nearly 8 times faster summed area table (4096×4096 image) in 20 lines of code including a fully customized CUDA schedule, as compared to 10 lines in NVIDIA Thrust which does not allow tiling or customizing the CUDA schedule.

This ongoing work is a collaboration with Jonathan Ragan-Kelley (Stanford University), Sylvain Paris (Adobe) and Fredo Durand (MIT).

6.2.3. *Video based rendering*

Participants: Abdelaziz Djelouah, George Drettakis.

In this project our objective is to propose a new algorithm for novel view synthesis in the case of dynamic scene. The main difference compared to static image-based rendering is the limited number of viewpoints and the presence of the extra time dimension. In a configuration where the number of cameras is limited, segmentation becomes crucial to identify moving foreground regions. To facilitate the difficult task of multi-view segmentation, we currently target scenes captured with stereo cameras. Stereo pairs provide important information on the geometry of the scene while simplifying the segmentation problem.

This ongoing work is a collaboration with Gabriel Brostow from University College London in the context of the CR-PLAY EU project.

6.2.4. *Temporally Coherent Video De-Anaglyph*

Participants: Joan Sol Roo, Christian Richardt.

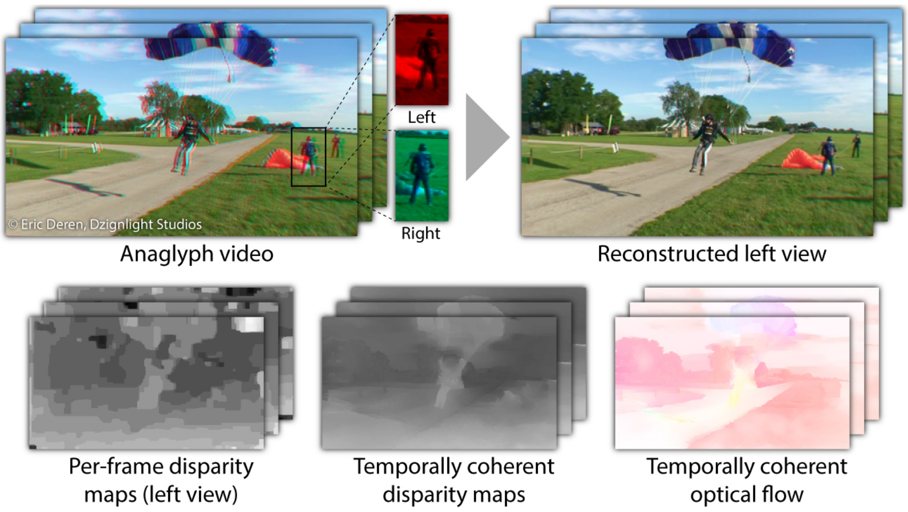


Figure 4. Top: We convert anaglyph videos (left) to temporally coherent full-color stereo videos (right). Bottom: Our approach starts with rough, per-frame disparity maps (left) and produces temporally coherent disparity maps and optical flow (center and right) that are used for reconstructing the stereo views.

This work investigates how to convert existing anaglyph videos to the full-color stereo format used by modern displays. Anaglyph videos only contain half the color information compared to the full-color videos, and the missing color channels need to be reconstructed from the existing ones in a plausible and temporally coherent fashion. In our approach, we put the temporal coherence of the stereo video results front and center (see Figure 4). As a result, our approach is both efficient and temporally coherent. In addition, it computes temporally coherent optical flow and disparity maps that can be used for various post-processing tasks. As a practical contribution, we also make the source code of our implementation available online under CeCILL-B license.

This work was carried out by Joan Sol Roo during his internship in the summer of 2013. The work was presented as a talk and poster at SIGGRAPH 2014 [19].

6.2.5. Probabilistic Connection Path Tracing

Participants: George Drettakis, Stefan Popov.

Bi-directional path tracing (BDPT) with Multiple Importance Sampling (MIS) is one of the most versatile unbiased rendering algorithms today. BDPT repeatedly generates sub-paths from the eye and the lights, which are connected for each pixel and then discarded. Unfortunately, many such bidirectional connections turn out to have low contribution to the solution. The key observation in this project, is that we can find better connections to an eye sub-path by considering multiple light sub-paths at once and creating connections probabilistically only with the most promising ones. We do this by storing light paths, and estimating probability density functions (PDF) of the discrete set of possible connections to all light paths. This has two key advantages: we efficiently create connections with high quality contributions by Monte Carlo sampling, and we reuse light paths across different eye paths. We also introduce a caching scheme for PDFs by deriving a low-dimensional approximation to sub-path contribution.

This ongoing work is a collaboration with Fredo Durand from MIT and Ravi Ramamoorthi from the University of California San Diego in the context of the CRISP associate team.

6.2.6. Unified Color and Texture Transfer for By-Example Scene Editing

Participants: Fumio Okura, Kenneth Vanhoey, Adrien Bousseau, George Drettakis.

Color and texture transfer methods are at the heart of by-example image editing techniques. Color transfer well represents the change of overall scene appearance; however it does not represent the change of texture and shape. On the other hand, by-example texture transfer expresses the texture change but it often destroys the target scene structure. We seek the best combination of by-example color and texture transfer to combine these transfer methods so as to selectively work where each method is suitable. Given the source and exemplar pair, the proposed algorithm learns local error metrics which describe if local change between the source and exemplar is best expressed by color or texture transfer. The metric provides us with a local prediction of where we need to synthesize textures using a texture transfer method. This work is a collaboration with Alexei Efros from UC Berkeley in the context of the associate team CRISP.

6.2.7. Improved Image-Based Rendering

Participants: Rodrigo Ortiz Cayon, Abdelaziz Djelouah, George Drettakis.

Image-based rendering algorithms based on warping present strong artifacts when rendering surfaces at grazing angle. We are working on a new IBR algorithm that overcomes this problem by rendering superpixel segments as piece-wise homography transformations. The input to our method is a set of images calibrated and a 3D point cloud generated from multi-view stereo reconstruction. In pre-processing we robustly fit planes to superpixel segments that contain reconstruction information and then propagate plausible depth and normal information for image-based rendering. Novel views are obtained by re-projecting superpixel segments as homography from different input views, then adaptively blending them according to distortion and confidence estimations.

6.2.8. Structured Procedural Textures

Participants: Kenneth Vanhoey, George Drettakis.

Textures form a popular tool to add visual detail to shapes, objects and scenes. Manual texture design is however a time-consuming process. An alternative is to generate textures from an input exemplar (*i.e.* an acquired photograph) automatically. The difficulty is to synthesize textures of arbitrary size from a single input, preferably with no repetition artifacts. State of the art synthesis techniques can be categorized in two: copy-based techniques and procedural noise-based ones. The first copy pixels using iterative algorithms. The latter deduce a continuous mathematical function from the exemplar, and evaluate it on the space to be textured. They have the advantage of continuity (no resolution-dependence, minimized memory storage, etc.) and fast local evaluation suitable for parallel GPU implementation. They are however tedious to define and manipulate. Current state of the art methods are limited to reproducing Gaussian patterns, that is, textures with no or few structure.

We investigate how to go beyond this limit. Noise-based methods constrain the Fourier power spectrum of a texture-generating noise function to resemble the spectrum of the exemplar. By also constraining the phase of the Fourier spectrum to resemble the exemplar, an exact reproduction is obtained, thus lacking variety and showing maximal repetition. By randomizing the phases, an unstructured "same-looking" image is obtained. This is suitable for noise-like patterns (*e.g.*, marble, wood veins, sand) but not for structured ones (*e.g.*, brick wall, mountain rocks, woven yarn).

In this project, we proceed by investigating the phase spectrum of an image. It contains the structure but identifying how and where is difficult. To characterize structure, we will exploit the splatting process of local random-phase noise and exhibit possible correlations between local phases and spatial placement.

This ongoing work is a collaboration with Ian Jermyn from Durham University.

6.3. Perception for Plausible Rendering

6.3.1. An Automated High Level Saliency Predictor for Smart Game Balancing

Participant: George Drettakis.

Successfully predicting visual attention can significantly improve many aspects of computer graphics: scene design, interactivity and rendering. Most previous attention models are mainly based on low-level image features, and fail to take into account high level factors such as scene context, topology, or task. Low-level saliency has previously been combined with task maps, but only for predetermined tasks. Thus, the application of these methods to graphics (e.g., for selective rendering) has not achieved its full potential.

In this work, we present the first automated high-level saliency predictor incorporating two hypotheses from perception and cognitive science that can be adapted to different tasks. The first states that a scene is comprised of objects expected to be found in a specific context as well objects out of context which are salient (scene schemata) while the other claims that viewer's attention is captured by isolated objects (singletons). We proposed a new model of attention by extending Eckstein's Differential Weighting Model. We conducted a formal eye-tracking experiment which confirmed that object saliency guides attention to specific objects in a game scene and determined appropriate parameters for a model. We presented a GPU-based system architecture that estimates the probabilities of objects to be attended in real-time (Figure 5). We embedded this tool in a game level editor to automatically adjust game level difficulty based on object saliency, offering a novel way to facilitate game design. We perform a study confirming that game level completion time depends on object topology as predicted by our system.

This work is a collaboration with George Alex Koulieris and Katerina Mania from the Technical University of Crete and Douglas Cunningham from the Technical University of Cottbus. The work was published in the ACM Transactions on Applied Perception (TAP) Journal [15] and presented as a Talk at SIGGRAPH 2014 in Vancouver.

6.3.2. C-LOD: Context-aware Material Level-of-Detail applied to Mobile Graphics

Participant: George Drettakis.

Attention-based Level-Of-Detail (LOD) managers downgrade the quality of areas that are expected to go unnoticed by an observer to economize on computational resources. The perceptibility of lowered visual fidelity is determined by the accuracy of the attention model that assigns quality levels. Most previous attention based LOD managers do not take into account saliency provoked by context, failing to provide consistently accurate attention predictions.

Figure 5. A low level saliency algorithm indicates that the most salient area of the image is the dark area behind the chair. Our tool highlights the vase at a consistent/singleton location as the most salient object in the image.

In this work, we extended a recent high level saliency model with four additional components yielding more accurate predictions: an object-intrinsic factor accounting for canonical form of objects, an object-context factor for contextual isolation of objects, a feature uniqueness term that accounts for the number of salient features in an image, and a temporal context that generates recurring fixations for objects inconsistent with the context. We conducted a perceptual experiment to acquire the weighting factors to initialize our model. We then designed C-LOD, a LOD manager that maintains a constant frame rate on mobile devices by dynamically re-adjusting material quality on secondary visual features of non-attended objects. In a proof of concept study we established that by incorporating C-LOD, complex effects such as parallax occlusion mapping usually omitted in mobile devices can now be employed, without overloading GPU capability and, at the same time, conserving battery power. We validated our work via eye-tracking (Figure 6)

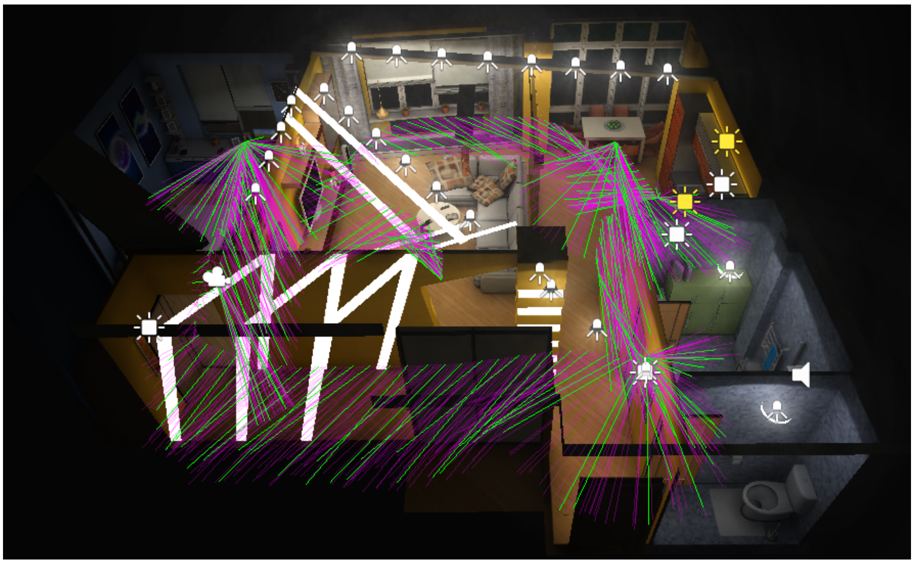


Figure 6. Our validation tool indicates the subject's gaze point with magenta colored beams. The green beams indicate predictions by our attention model.

This work is a collaboration with George Alex Koulieris and Katerina Mania from the Technical University of Crete and Douglas Cunningham from the Technical University of Cottbus. The work was published in a special issue of Computer Graphics Forum [] and was presented at the Eurographics Symposium on Rendering 2014 in Lyon. It was also presented as a poster at SIGGRAPH 2014 in Vancouver winning the 3rd place in the ACM's Graduate Student Research Competition.

6.4. Interaction and Design for Virtual Environments

6.4.1. Evaluation of Direct Manipulation using Finger Tracking for Complex Tasks in an Immersive Cube

Participants: Emmanuelle Chapoulie, George Drettakis.

We present a solution for interaction using finger tracking in a cubic immersive virtual reality system (or immersive cube). Rather than using a traditional flystick device, users can manipulate objects with fingers of both hands in a close-to-natural manner for moderately complex, general purpose tasks. Our solution couples finger tracking with a real-time physics engine, combined with a heuristic approach for hand manipulation, which is robust to tracker noise and simulation instabilities. We performed a first study to evaluate our interface with tasks involving complex manipulations, such as balancing objects while walking in the cube. The users finger-tracked manipulation was compared to manipulation with a 6 degree-of-freedom flystick, as well as with carrying out the same task in the real world. Users were also asked to perform a free task, allowing us to observe their perceived level of presence in the scene. Our results showed that our approach provides a feasible interface for immersive cube environments and is perceived by users as being closer to the real experience compared to the flystick. However, the flystick outperforms direct manipulation in terms of speed and precision.

This work is a collaboration with Maria Roussou and Evanthia Dimara from the University of Athens, Maud Marchal from Inria Rennes, and Jean-Christophe Lombardo from Inria Sophia Antipolis. The work has been published in the journal Virtual Reality [13].

Figure 7. A user balancing a tray with both hands [13].

We have also worked on a followup study in which we examine a much more controlled context, studying only very limited movements, in 1D, 2D and 3D. To do this we designed specific devices that can be instantiated both in the virtual world and as physical objects. We compared finger manipulation to wand and to real configurations; the study demonstrated the feasibility of such a controlled comparison for the study of finger-based interaction. This work is in collaboration with InSitu, specifically F. Tsandilas, W. Mackay and L. Oehlberg, and has been accepted for publication in 2015 at IEEE 3DUI.

6.4.2. Reminiscence Therapy using Image-Based Rendering in VR

Participants: Emmanuelle Chapoulie, George Drettakis, Rachid Guerchouche, Gaurav Chaurasia.

We present a novel VR solution for Reminiscence Therapy (RT), developed jointly by a group of memory clinicians and computer scientists. RT involves the discussion of past activities, events or experiences with others, often with the aid of tangible props which are familiar items from the past; it is a popular intervention in dementia care. We introduced an immersive VR system designed for RT, which allows easy presentation of familiar environments. In particular, our system supports highly-realistic Image-Based Rendering in an immersive setting. To evaluate the effectiveness and utility of our system for RT, we performed a study with healthy elderly participants to test if our VR system could help with the generation of autobiographical memories. We adapted a verbal Autobiographical Fluency protocol to our VR context, in which elderly participants were asked to generate memories based on images they were shown. We compared the use of our image-based system for an unknown and a familiar environment. The results of our study showed that the number of memories generated for a familiar environment is higher than the number of memories obtained for an unknown environment using our system. This indicates that IBR can convey familiarity of a given scene, which is an essential requirement for the use of VR in RT. Our results also showed that our system is as effective as traditional RT protocols, while acceptability and motivation scores demonstrated that our system is well tolerated by elderly participants.

This work is a collaboration with Pierre-David Petit and Philippe Robert from the CMRR in Nice. The work has been published in the Proceedings of IEEE Virtual Reality [18].

Figure 8. Left: our hardware setup. Right: new point of view reconstructed from input cameras.

6.4.3. Lightfield Editing

Participant: Adrien Bousseau.

Lightfields capture multiple nearby views of a scene and are consolidating themselves as the successors of conventional photographs. As the field grows and evolves, the need for tools to process and manipulate lightfields arises. However, traditional image manipulation software such as Adobe Photoshop are designed to handle single views and their interfaces cannot cope with multiple views coherently. We conducted a thorough study to evaluate different lightfield editing interfaces, tools and workflows from a user perspective. We additionally investigate the potential benefits of using depth information when editing, and the limitations imposed by imperfect depth reconstruction using current techniques. We perform two different experiments, collecting both objective and subjective data from a varied number of point-based editing tasks of increasing complexity: In the first experiment, we rely on perfect depth from synthetic lightfields, and focus on simple edits. This allows us to gain basic insight on lightfield editing, and to design a more advanced editing interface. This is then used in the second experiment, employing real lightfields with imperfect reconstructed depth, and covering more advanced editing tasks. Our study shows that users can edit lightfields with our tested interface and tools, even in the presence of imperfect depth. They follow different workflows depending on the task at hand, mostly relying on a combination of different depth cues. Last, we confirm our findings by asking a set of artists to freely edit both real and synthetic lightfields.

This work is a collaboration with Adrian Jarabo, Belen Masia and Diego Gutierrez from Universidad de Zaragoza and Fabio Pellacini from Sapienza Universita di Roma. This work was published at ACM Transactions on Graphics 2014 (Proc. SIGGRAPH) [14].

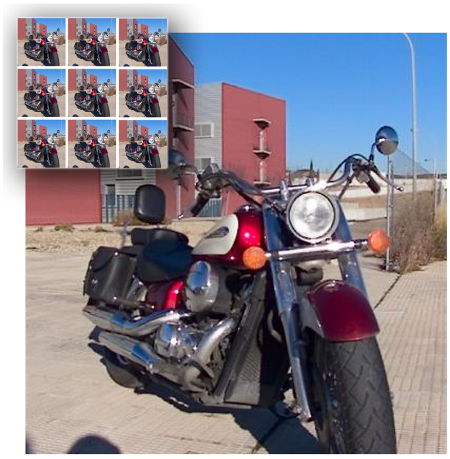


Figure 9. A lightfields represents multiple nearby views of a scene. We conducted a study to evaluate how people edit such data.

6.4.4. True2Form: 3D Curve Networks from 2D Sketches via Selective Regularization

Participant: Adrien Bousseau.

True2Form is a sketch-based modeling system that reconstructs 3D curves from typical design sketches. Our approach to infer 3D form from 2D drawings is a novel mathematical framework of insights derived from perception and design literature. We note that designers favor viewpoints that maximally reveal 3D shape information, and strategically sketch descriptive curves that convey intrinsic shape properties, such as curvature, symmetry, or parallelism. Studies indicate that viewers apply these properties selectively to envision a globally consistent 3D shape. We mimic this selective regularization algorithmically, by progressively detecting and enforcing applicable properties, accounting for their global impact on an evolving 3D curve network. Balancing regularity enforcement against sketch fidelity at each step allows us to correct for inaccuracy inherent in free-hand sketching. We perceptually validate our approach by showing agreement between our algorithm and viewers in selecting applicable regularities. We further evaluate our solution by: reconstructing a range of 3D models from diversely sourced sketches; comparisons to prior art; and visual comparison to both ground-truth and 3D reconstructions by designers.

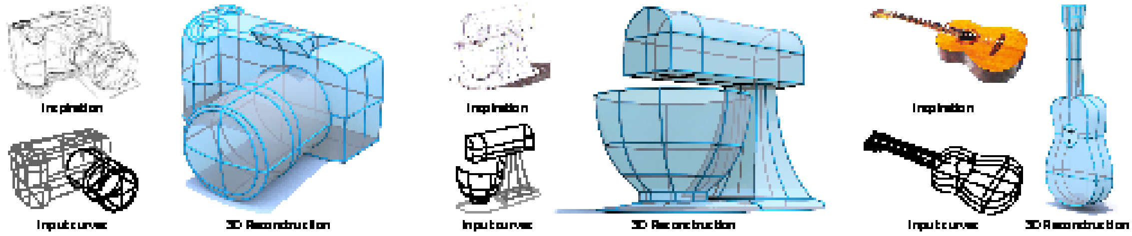


Figure 10. Our single-view modeling system allows us to reconstruct 3D models by tracing curves over existing sketches and photographs.

This work is a collaboration with James McCrae and Karan Singh from the University of Toronto and Xu Baoxuan, Will Chang and Alla Sheffer from the University of British Columbia. The paper was published at ACM Transactions on Graphics 2014 (Proc. SIGGRAPH) [17].

6.4.5. BendFields: Regularized Curvature Fields from Rough Concept Sketches

Participants: Adrien Bousseau, Emmanuel Iarussi.

Designers frequently draw curvature lines to convey bending of smooth surfaces in concept sketches. We present a method to extrapolate curvature lines in a rough concept sketch, recovering the intended 3D curvature field and surface normal at each pixel of the sketch. This 3D information allows us to enrich the sketch with 3D-looking shading and texturing. We first introduce the concept of *regularized curvature lines* that model the lines designers draw over curved surfaces, encompassing curvature lines and their extension as geodesics over flat or umbilical regions. We build on this concept to define the orthogonal cross field that assigns two regularized curvature lines to each point of a 3D surface. Our algorithm first estimates the projection of this cross field in the drawing, which is non-orthogonal due to foreshortening. We formulate this estimation as a scattered interpolation of the strokes drawn in the sketch, which makes our method robust to sketchy lines that are typical for design sketches. Our interpolation relies on a novel smoothness energy that we derive from our definition of regularized curvature lines. Optimizing this energy subject to the stroke constraints produces a dense non-orthogonal 2D cross field, which we then lift to 3D by imposing orthogonality. Thus, one central concept of our approach is the generalization of existing cross field algorithms to the non-orthogonal case. We demonstrate our algorithm on a variety of concept sketches with various levels of sketchiness. We also compare our approach with existing work that takes clean vector drawings as input.

This work is a collaboration with David Bommes from Titane project team, Inria Sophia-Antipolis. The manuscript has been accepted for publication with minor revisions at ACM Transactions on Graphics (TOG).

6.4.6. Line Drawing Interpretation in a Multi-View Context

Participant: Adrien Bousseau.

Many design tasks involve the creation of new objects in the context of an existing scene. Existing work in computer vision only provides partial support for such tasks. On the one hand, multi-view stereo algorithms allow the reconstruction of real-world scenes, while on the other hand algorithms for line-drawing interpretation do not take context into account. This work combines the strength of these two domains to interpret line drawings of imaginary objects drawn over photographs of an existing scene. The main challenge we face is to identify the existing 3D structure that correlates with the line drawing while also allowing the creation of new structure that is not present in the real world. We propose a labeling algorithm to tackle this problem, where some of the labels capture dominant orientations of the real scene while a free label allows the discovery of new orientations in the imaginary scene.

This work is a collaboration with Jean-Dominique Favreau and Florent Lafarge from Titane project team, Inria Sophia-Antipolis and is under submission for the CVPR conference.

6.4.7. Wrap It! Computer-Assisted Design and Fabrication of Wire Wrapped Jewelry

Participants: Adrien Bousseau, Emmanuel Iarussi.

We developed an interactive tool to assist the process of creating and crafting wire wrapped pieces of jewelry. In a first step, we guide the user in conceiving designs which are suitable to be fabricated with metal wire. In a second step, we assist fabrication by taking inspiration from jigs-based techniques, frequently used by craftsmen as a way to guide and support the wrapping process. Given a vector drawing composed of curves to be fabricated, it is crucial to first decompose it into segments that can be constructed with metal wire. Literature on jewelry-making provides a wide range of examples to perform this task, but they are hard to generalize to any input design. Based on the observation of these examples, we distill and generalize a set of design principles behind the finished pieces of jewelry. Relying on those principles, we propose an algorithm that generates a decomposition of the input where each piece is a single component of wire, that can be wrapped and gathered with the others. In addition, we also automate the design of custom physical jigs for fabrication of the jewelry piece. A jig consists of a board with holes on it, arranged in a regular grid structure. By placing a set of pins (of different radius) on the jig, the craftman builds a support structure that guides the wrapping process. The wire is bended and twisted around those pins to create the shape. Given the input design curves and the available jig parameters (size, number and radius of the pins), we propose an algorithm to automatically generate an arrangement of pins in order to better approximate the input curve with wire. Finally, users can follow automatically-generated step-by-step instructions to place the pins in the jig board and fabricate the end piece of jewelry.

This ongoing work is a collaboration with Wilmot Li from Adobe, San Francisco. The project was initiated by a 3-months visit of Emmanuel Iarussi at Adobe.

6.4.8. Studying how novice designers communicate with sketches and prototypes

Participant: Adrien Bousseau.

We performed a user study to better understand how novice designers communicate a concept during the different phases of its development. Our study was conducted as a one-day design contest where participants had to propose a concept, present it to a jury, describe it to an engineer and finally fabricate a prototype with the help of another participant. We collected sketches and videos for all steps of this exercise in order to evaluate how the concept evolves and how it is described to different audiences. We hope that our findings will inform the development of better computer-assisted design tools for novices.

This is an ongoing work in collaboration with Wendy McKay, Theophanis Tsandilas and Lora Oehlberg from the InSitu project team - Inria Saclay, in the context of the ANR DRAO project.

6.4.9. Vectorising Bitmaps into Semi-Transparent Gradient Layers

Participants: Christian Richardt, Adrien Bousseau, George Drettakis.

Vector artists create complex artworks by stacking simple layers. We demonstrate the benefit of this strategy for image vectorisation, and present an interactive approach for decompositing bitmap drawings and studio photographs into opaque and semi-transparent vector layers. Semi-transparent layers are especially challenging to extract, since they require the inversion of the non-linear compositing equation. We make this problem tractable by exploiting the parametric nature of vector gradients, jointly separating and vectorising semi-transparent regions. Specifically, we constrain the foreground colours to vary according to linear or radial parametric gradients, restricting the number of unknowns and allowing our system to efficiently solve for an editable semi-transparent foreground.

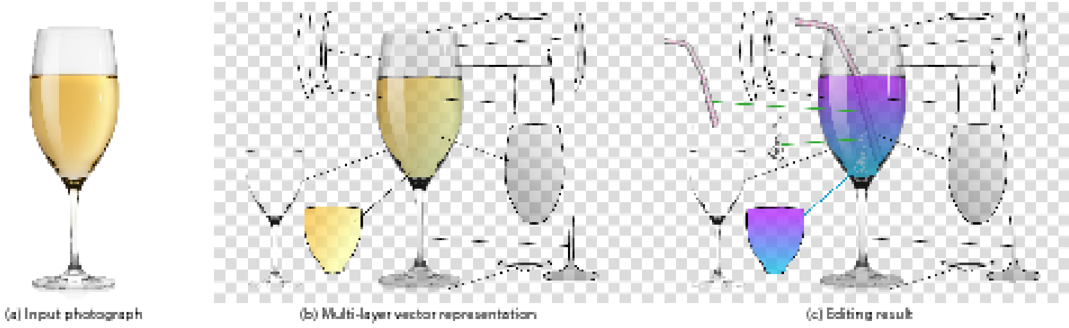


Figure 11. Our interactive vectorisation technique lets users vectorise an input bitmap (a) into a stack of opaque and semi-transparent vector layers composed of linear or radial colour gradients (b). Users can manipulate the resulting layers using standard tools to quickly produce new looks (c). Semi-transparent layers are outlined for visualisation; these edges are not part of our result.

We propose a progressive workflow, where the user successively selects a semi-transparent or opaque region in the bitmap, which our algorithm separates as a foreground vector gradient and a background bitmap layer. The user can choose to decompose the background further or vectorise it as an opaque layer. The resulting layered vector representation allows a variety of edits, as illustrated in Figure 11 , such as modifying the shape of highlights, adding texture to an object or changing its diffuse colour. Our approach facilitates the creation of such layered vector graphics from bitmaps, and we thus see our method as a valuable tool for professional artists and novice users alike.

This work is a collaboration with Jorge Lopez-Moreno, now a postdoc at the University of Madrid, and Maneesh Agrawala from the University of California, Berkeley in the context of the CRISP Associated Team. The paper was presented at the Eurographics Symposium on Rendering (EGSR) 2014, and is published in a special issue of the journal Computer Graphics Forum [16].

RITS Team

6. New Results

6.1. Highlights of the Year

YoGoKo [0], a startup company of RITS, was founded in 2014 by employees from three research institutes: Mines ParisTech, Telecom Bretagne and Inria. YoGoKo makes use of softwares developed in teams specialized in Internet technologies. RSM (Telecom Bretagne), CAOR (Mines ParisTech) and RITS (Inria) are research teams have been working together since 2006 on innovative communication solutions applied to Intelligent Transportation Systems. They contributed to several collaborative R& D projects related to ITS (CVIS, ITSSv6, GeoNet, DriveC2X, SCORE@F, . . .). In 2012, these laboratories engaged together into the development of a common demonstration platform which comprises connected vehicles (fleet of conventional vehicles from Mines ParisTech and fleet of autonomous vehicles from Inria), roadside equipments and cloud-based services. YoGoKo demonstration platform was finally revealed on Feb. 11 th 2014 during the Mobility2.0 event organized by the French Ministry of Transport. This successful demonstation and the extremely warmfull feedack gained at this occasion triggered the launch of YoGoKo as a company. YoGoKo develops innovative communication solutions for fixed and mobile multi-connected devices. The objective is to maintain secure and continuous connectivity with their communication peers, either in their immediate environment or a remote location (control centers or Internet hosts).

6.2. Development of a Platform for Arbitration and Sharing Control Applications

Participants: David Gonzalez Bautista, Vicente Milanes Montero, Fawzi Nashashibi, Joshué Pérez Rastelli.

RITS have been leading the activities in the framework DESERVE project, related to arbitration and control sharing in automated vehicle. The analysis of existing vehicle control (and arbitration) solutions, considering the driver in the control loop is the main challenge of this work. We consider sharing control techniques and different solutions in the task management. New standard in the taxonomy of autonomous driving, as the SAE J3016, are considered in the arbitration and sharing control design. The aim is to allow the applications to make effective use of the driver model to improve the acceptability of the functions developed, as: Driver Drowsiness and Driver intention.

The arbitration module is defined into the IWI manager of the DESERVE abstraction. This component determines the action to be taken by the driver. The Driver Assistance Systems involve two main decision makers: the driver and the automated systems. This module considers different inputs, as follow: the Trajectory planning, Driver stage, Risk Management. The output determines who should take the control of the vehicle and the level of arbitration (or disposal) of the driver in different situations. This work uses the software tool FEMOT (Fuzzy Embedded MOTor). More detail can be found in [46].

6.3. Optimal Energy Consumption for Urban Electric Vehicles

Participants: David Gonzalez Bautista, Vicente Milanes Montero, Joshué Pérez Rastelli.

[0] http://www.yogoko.fr/

RITS team is specially supporting two kinds of transport systems: electric mass-produced vehicles and Cybernetic Transport Systems (also electrically propelled) for urban environments. One of the key factors for getting a higher market penetration of such vehicles is their autonomy. Having this in mind, the goal of this research line is to create optimal algorithms for improving electric vehicles' battery life. It covers two specific arenas: 1) determining optimal path planning in terms of energy saving (proposed for 2015); and 2) once the route is determined, generating an adequate speed profile for covering that path. The latter objective has been investigated during 2014. Energetic model of vehicle dynamics have been developed in order to determine the lowest consumption for each of the route segments. It has permitted to develop speed references between segments combined with polynomial transition functions for the whole route to be covered. Additionally, a high-level fuzzy controller has been also designed to make the system robust to low-level failures on reference tracking. Up to 20% of battery savings have been obtained in the first tests with the proposed algorithm, showing the proper performance of the system. Additional work for adding more information from the environment as other road agents or potential unexpected diversion on the road will be also investigated during 2015 for adapting the algorithm to more realistic environments. This work has been also developed in cooperation with MSc students from Simon Bolivar University (Venezuela) and AGMUS University System (Puerto Rico, US).

6.4. Perception and control strategies for autonomous docking for electric freight vehicles

Participants: Joshué Pérez Rastelli, Evangeline Pollard, Vicente Milanes Montero, Fawzi Nashashibi.

The freight transportation is defined as the process of carrying goods and persons from one given point to another. Recently, urban freight transportations have been used as an alternative for the delivery problems of goods in urban environments. The present work is developed in the framework of the Furbot project (FP7), which presents a solution for future urban freight transport with new light-duty architecture with full-electrical vehicles. We focused on the onboard intelligent units, dedicated to improve the perception and control systems onboard the vehicle for the parking/docking process, considering loading and unloading phases of the freight transport procedure. Two lasers were placed on the vehicle in order to localize it with respect to the freight box. A polynomial approach is used for the trajectory planning for a smooth docking maneuver. This proposal was first tested in a 3D simulator, and then validated in a real platform. The results presented in [45] shows the good behavior of our approach, which will be implemented in the FURBOT vehicle at the end of the project.

6.5. Description and technical specification of Cybernetic Transportation Systems: an urban transportation concept

Participants: Joshué Pérez Rastelli, Vicente Milanes Montero, David Gonzalez Bautista, Armand Yvet, Fawzi Nashashibi.

The Cybernetic Transportation Systems (CTS) is an urban transportation concept based on two ideas: the car sharing and the automation of dedicated systems with door-to-door capabilities. In the last decade, many European projects have been developed in this context, where some of the most important are: Cybercars, Cybercars II, CyberMove, CyberC3 and CityMobil, where a first fleet of vehicles were developed by different companies and research centers around Europe, Asia and America. Considering these previous works, the FP7 project Citymobil II is in progress since 2012. Its goal is to solve some of the limitations found so far, including the definition of the legal framework for autonomous vehicles on urban environment. Much of the perception and control software has been improved in the Inria's Cybus. New guidance functionalities were developed, mainly with the introduction of stereovision-based SLAM, and Bezier curve in path planning generation. In this work, automated CTSs involved are used in the different showcases in European cities. This work presents the different improvements, adaptation and instrumentation of the vehicle used. Results show tests in our facilities at Inria-Rocquencourt (France) and the first showcase at León (Spain).

6.6. Evidential Simultaneous Localization And Mapping to describe intersection

Participants: Guillaume Trehard, Evangeline Pollard, Fawzi Nashashibi.

Intersections management remains a tough challenge to tackle before reaching autonomous driving in urban environment. The field of view of the vehicle is often limited by several sensors occlusions, the shapes and priority rules can significantly differ from an intersection to another and road users from pedestrian to public transports have to cross each other in sometimes complex manners.

In this context, mapping the surrounding of the vehicle and being able to estimate its position regarding a global database is crucial.

A solution of *Simultaneous Localization And Mapping* (SLAM) have then been proposed based on a 2D LIDAR sensor [49]. In the rich SLAM literature, the originality of this method lays in the use of Transferable Belief Model (TBM) framework instead of a classic probabilistic one. If this proposition was just a change of mathematical context, TBM led to an explicit management of not-known and conflict information so that its application to SLAM algorithm appeared to be really effective and robust in crowded situations. The proposed solution indeed enables to provide a map of the *static* environment crossed by the vehicle and to detect mobile obstacles in the same process and without additional tracking system.

This *Evidential SLAM* have then been tested with success over different sequences and laser set-ups extracted from the KITTI database [50].

Researches are now focused on the fusion between this SLAM solution and a Global Navigation Satellite System (GNSS) receiver to enable a map-matching on a database such as Open Street Map.

6.7. Laser based road obstacle classification

Participants: Pierre Merdrignac, Evangeline Pollard, Oyunchimeg Shagdar, Fawzi Nashashibi.

Vehicle and pedestrian collisions often result in fatality to the vulnerable road users, indicating a strong need of technologies to protect such vulnerable road users. Laser sensors have been extensively used for moving obstacles detection and tracking. Laser impacts are produced by reflection on these obstacles which suggest that more information is available for their classification. This year, we introduced the design of a new system for road obstacles classification that is divided in four parts: definition of geometric features, selection of the best features, multi-class *segment* classification based on Support Vector Machines (SVM) and *track* classification from SVM decision values integration. Our study discloses a sorted list of useful features for road obstacle recognition that were used to construct a multi-class SVM. Finally, we tested our system with 2D and 3D laser sequences and shown that it can successfully estimate the class of some road obstacles around the vehicle.

6.8. Deformable Parts Model based approach for on-road object detection and classification

Participants: Wei-Lin Ku, Evangeline Pollard, Anne Verroust-Blondet.

An important perception problem for driver assistance is the detection of the road obstacles and the recognition of their type (cars, cycles, pedestrians). This year, we tackled the on-road objects detection problem by testing and improving vision-based methods. We proposed and compared several DPM based strategies for on-road object detection and classification, laying emphasis on the problem of detecting smaller/occluded cars and pedestrians. A hybrid approach combining detection from small/large models trained with different clustering method has been introduced to boost the detection performance in both Average Precision and Maximum Recall in every difficulty level. Finally, a geometry reasoning based filtering has been employed to eliminate false alarms while preserving a great deal amount of true positives. Experimental results showed the improvement both in hybrid and geometry reasoning approaches. Most of this work has been done during the internship of Wei-Lin Ku.

6.9. Saturated Feedback Control for an Automated Parallel Parking Assist System

Participants: Mohamed Marouf, Fawzi Nashashibi, Plamen Petrov.

In 2014, RITS extended its activities in the design and development of specific automated maneuvres. One particular interesting topic is the parallel parking problem of automatic front-wheel steering vehicles. The problem of stabilizing the vehicle at desired position and orientation is seen as an extension of the tracking problem. A saturated control has been proposed which achieves quick steering of the system near the desired position of the parking spot with desired orientation and can be successfully used in solving parking problems. In addition, in order to obtain larger area of the starting positions of the vehicle with respect to the parking spot for the first reverse maneuver of the parallel parking, an approach using saturated control with two different levels of saturation is proposed. The vehicle can be automatically parked by using one or multiple maneuvers, depending on the size of the parking spot. Simulation results were presented first in [44] to confirm the effectiveness of the proposed control schemes. New results extended to all types of parking lots shapes were recently obtained using this approach. The validation has been performed with real vehicles in the Inria test site.

6.10. Vehicle to pedestrian communications

Participants: Pierre Merdrignac, Oyunchimeg Shagdar, Evangeline Pollard, Fawzi Nashashibi.

Vehicle and pedestrian collisions often result in fatality and serious injury to the vulnerable road users. While vehicle to vehicle (V2V) communications have taken much attention in the academic and industrial sectors, very limited effort has been made for vehicle to pedestrian communications. Unlike the V2V cases, where antennas are often installed on the vehicle rooftop, pedestrian's handheld device can be carried in such a way e.g, in a bag or in a pocket, which results in poor and unpredictable communications quality. In this work, we seek to an answer to the question whether the Wi-Fi-based V2P communications meet the requirements of the pedestrian safety. This year, we studied the performances of the V2P communications especially for the receive signal strength, packet inter-arrival time, and message delivery ratio. Moreover, in order to demonstrate the feasibility of pedestrian safety supported by the V2P communications, we developed a software tool, V2ProVu, which has the functionalities of Wi-Fi based V2P communications, collision risk calculations, and hazard alarming. This work has been published in [34].

6.11. Multicast Communications for Cooperative Vehicular Systems

Participants: Ines Ben Jemaa, Oyunchimeg Shagdar, Arnaud de La Fortelle.

Vehicular communications allow emerging new multicast applications such as fleet management and point of interest (POI). Both applications require Internet-to-vehicle multicasting. These approaches could not be applied to vehicular networks (VANET) due to their dynamic and distributed nature. In order to enable such multicasting, our work deals with two aspects. First, reachability of the moving vehicles to the multicast service and second, multicast message dissemination in VANET. We propose a simplified approach that extends Mobile IP and Proxy Mobile IP. This approach aims at optimizing message exchange between vehicles and entities responsible for managing their mobility in Internet. To study the dissemination mechanisms that are suitable for fleet management applications, we propose to revisit traditional multicast routing techniques that rely on a tree structure. For this purpose, we study their application to vehicular networks. In particular, as vehicular networks are known to have changing topology, we study the application of Multicast Adhoc On Demand Vector, MAODV. We propose then Motion-MAODV [35] [16] an improved version of MAODV that aims at enhancing routes built by MAODV in vehicular networks and guarantee longer route lifetime. Finally, to enable geographic dissemination as required by POI applications, we propose the routing protocol Melody that provides a geocast dissemination in urban environments. Through simulations, Melody ensures more reliable and efficient packet delivery to a given geographic area compared to traditional geo-broadcasting schemes in highly dense scenarios.

6.12. Visible Light Communication for ITS applications

Participants: Mohammad Abu Alhoul, Oyunchimeg Shagdar, Fawzi Nashashibi.

Visible Light Communication (VLC) technology is an efficient supportive communication technology for platooning applications over short inter-vehicle distances. After implementing a complete VLC channel model, which enabling precise calculations of the optical link performance for different intervehicle distances presented in our previous work [1], this year we have studied and proposed tracking-alike method aiming at ensuring the continuity of the Line-of-Sight (LOS) and extending the Field of view (FOV) limitations. This method benefits from the exchanged information about the relative directional position of each member of the platoon, together with front and rear facing directions of each vehicle, which can be very useful data for building a reliable smooth geometrical-based compensation method. The simulation results showed that trajectory influences on the optical incidence and irradiance angles can be compensated efficiently and without deploying any tracking method.

6.13. Study on the IEEE 802.11p Channel Congestion Problem

Participant: Oyunchimeg Shagdar.

The IEEE 802.11p is a standardized WiFi technology dedicated to V2X communications for especially road safety and efficiency applications. It is expected that vehicles periodically broadcast messages to announce their existences using the IEEE 802.11p frequency channel. However, because the IEEE 802.11p has a limited wireless bandwidth, in dense traffic conditions the V2V communications performances are poor, failing to satisfy the application requirements. In RITS, we study the issue and develop congestion control algorithms. In 2014, we studied the reactive distributed congestion control algorithm, proposed by the European Telecommunications Standards Institute (ETSI), and showed that the algorithm creates unstable resource utilization, which can cause the reactive Distributed Congestion Control (DCC) to perform worse than non-DCC systems. We proposed an asynchronous algorithm, where DCC control is made in such a way that channel resource is used in an asynchronous manner by the different stations. Our results show that the asynchronous DCC approach outperforms both the non-DCC and reactive DCC mechanisms. The work has been reported at a ETSI meeting in December 18, 2014 [54].

6.14. Study on V2V Communications and Emergent Behavior of Heuristically-Driven Intelligent Vehicles

Participant: Oyunchimeg Shagdar.

The automated cooperative driving applications require efforts on multiple research domains including robotics, artificial intelligence, and communications to build a safe and intelligent collective driving behavior. While some studies show the potentials of the V2V communications for safer and smoother automated driving, it is still not clear if the standardized technologies can meet the strict requirements of the automated driving applications. More importantly, if the decisions for individual vehicles' control are based on the V2V communications, the communications performance must largely affect the "quality" of the collective behavior. Motivated by this, we study the inter-dependencies between communications and collective automated driving behavior. In our study [43] we combine different V2V communication modes with different dynamic path-finding heuristics, over a population of several hundreds of intelligent vehicles, to observe convergence towards stable traffic. The various traffic stability levels are compared in order to exhibit most efficient combinations of communication modes and path-finding heuristics.

6.15. Distributed Agreement and String Control in Intelligent Vehicular Networks (IVNs)

Participant: Gérard Le Lann.

IVNs are composed of automated (autonomous and communicating) vehicles, ranging from pre-planned platoons to ad hoc vehicular networks (VANETs). Agreement problems in the presence of concurrency and failures are not well investigated yet in IVNs. We have examined a specific class of such problems, those arising in string formations. Regarding string membership (vehicles leaving or joining a string), with few exceptions, safety issues have been addressed so far assuming that (1) no more than 1 insertion operation would be performed at any given time or, (2) every vehicle decides unilaterally, i.e. undertakes a maneuver after having activated some signal, leaving to surrounding vehicles the responsibility of inferring intended maneuvers. Assumption (1) is not realistic. There are numerous risk-prone scenarios where a posteriori reactive approaches (assumption (2)) may fail. Therefore the need for investigating proactive approaches, where vehicles (1) are made aware of intended impending maneuvers, (2) agree on which maneuvers can be safely undertaken, prior to performing physical maneuvers. It follows that a solution to numerous string control problems consists of a pair (A,Φ), where A stands for a distributed agreement algorithm which achieves global coordination in the presence of failures and concurrency, and Φ stands for control laws drawn from control theory and robotics. Algorithm A is run prior to local activations of Φ.
Our work is based on the cohort construct, which serves to formalize the concept of strings. Velocity Agreement is the generic problem selected. At any time, some number of string/cohort members may propose each a new velocity value. In fine, every vehicle computes a unique new velocity V. Proposed values are propagated via neighbor-to-neighbor (N2N) radio communications. We have devised a solution called the VAgree algorithm. In the presence of up to f failures (on-board systems, N2N message losses), the following properties shall hold:

- Validity: Decision value $V = \Psi$(proposed values).
- Agreement: No two members decide differently.
- Time-Bounded Termination: VAgree terminates at most θ time units.
- Synchronicity: Times at which V is posted to on-board systems are comprised within a small time interval ϵ. Distance traveled during ϵ by the member earliest to post V until the latest member does so is an order of magnitude smaller than vehicle sizes.

The VAgree algorithm is presented in a paper which is under submission.

6.16. Standardization and automated vehicles

Participant: Michel Parent.

Michel Parent has been active over the last 4 years in this group to introduce automated vehicles in the scope of service robots and he contributed actively in the activities of several working groups (WGs). In WG7, he participated in the writing of the document ISO13482 on the safety of service robots. This document has reached the Final Draft for an International Standard level (FDIS) and is now published in English and French. It has already been used by companies to certify some robotics products, including Robosoft in France for automated vehicles. In WG8, Michel Parent participated in the elaboration of standard procedures for the testing of service robots and in particular for automated vehicles. The document CD18646 « Robots and robotic devices — Performance criteria and related test methods for service robots — Part 1: Locomotion for wheeled robots » is in progress.

6.17. Legislation and certification of fully automated road vehicles for urban public transport

Participant: Michel Parent.

An important research area of automated road vehicles and one of the focuses of the CityMobil2 Project is to look at the legislation and certification of fully automated road vehicles for urban public transport (the cybercars). This part of the research was done essentially by Michel Parent in 2014 and gave birth to several CityMobil2 deliverables.

One of the tasks was to identify the current legislation in France and the organizations involved in the changes for this legislation. Several meetings were therefore organized at the French level with key persons from the Ministry of Transport, the Ministry of Interior (responsible for the road legislation) and their services (in particular the SRMTG in charge of certifying the guided transport systems). These meetings were essential in obtaining the authorization to operate the cybercars for the demonstration in La Rochelle. At the European level, a meeting was organized in May 2014 with representatives of 12 of the European countries (mostly those involved directly with CityMobil2 or with automated vehicles R&D).

Another task was to propose a certification methodology for automated road transport systems. For this task, a careful analysis of the test site in La Rochelle was conducted and led to a number of use cases. Key elements were defined to perform the risk analysis. Many hazards were identified but the most important ones are the behavior of pedestrian and cyclists. For the analysis, two key variables were defined: the minimum mobile object detection distance (MMODD) and the maximum mobile object speed (MMOS). For each use case, a combination of these 2 variable lead to a maximum vehicle speed in order to reach an acceptable risk evaluated as a combination of severity and probability.

In order to verify the proper behavior of the vehicle itself (lane tracking, obstacle avoidance, comfort,...), a number of standard tests have also been defined and are now proposed at the International level (ISO standards).

6.18. Belief propagation inference for traffic prediction

Participants: Cyril Furtlehner, Jean-Marc Lasgouttes.

This work [60] deals with real-time prediction of traffic conditions in a setting where the only available information is floating car data (FCD) sent by probe vehicles. The main focus is on finding a good way to encode some coarse information (typically whether traffic on a segment is fluid or congested), and to decode it in the form of real-time traffic reconstruction and prediction. Our approach relies in particular on the belief propagation algorithm.

These studies have been done in particular in the framework of the projects Travesti and Pumas.

This year, the work about the theoretical aspects of encoding real valued variables into a binary Ising model has been under review for a Journal and has been largely revised in the process.

6.19. Sparse covariance inverse estimate for Gaussian Markov Random Field

Participants: Cyril Furtlehner, Jean-Marc Lasgouttes.

We investigate the problem of Gaussian Markov random field selection under a non-analytic constraint: the estimated models must be compatible with a fast inference algorithm, namely the Gaussian belief propagation algorithm. To address this question, we introduce the ☆-IPS framework, based on iterative proportional scaling, which incrementally selects candidate links in a greedy manner. Besides its intrinsic sparsity-inducing ability, this algorithm is flexible enough to incorporate various spectral constraints, like e.g. walk summability, and topological constraints, like short loops avoidance. Experimental tests on various datasets, including traffic data from San Francisco Bay Area, indicate that this approach can deliver, with reasonable computational cost, a broad range of efficient inference models, which are not accessible through penalization with traditional sparsity-inducing norms.

This work has been presented at ECML/PKDD 2014 [40]. The code for ☆-IPS has been made available at https://who.rocq.inria.fr/Jean-Marc.Lasgouttes/star-ips/.

6.20. Herding behavior in a social game

Participants: Guy Fayolle, Jean-Marc Lasgouttes.

The system *Ma Micro Planète* belongs to the so-called *Massively Multi-Player online Role Playing game* (MMORPG), its main goal being to incite users to have a sustainable mobility. Two objectives have been pursued.

- Construct an experimental platform to collect data in order to prompt actors of the mobility to share information (open data system).
- See how various mechanisms of a game having an additive effect could modify the transportation requests.

At the heart of the game are community-driven *points of interest* (POIs), or *sites*, which have a score that depends on the players activity. The aim of this work is to understand the dynamics of the underlying stochastic process. We analyze in detail its stationary regime in the thermodynamic limit, when the number of players tends to infinity. In particular, for some classes of input sequences and selection policies, we provide necessary and sufficient conditions for the existence of a complete meanfield-like measure, showing off an interesting *condensation* phenomenon.

The work has been published this year in *Queueing Systems* [20].

6.21. Properties of random walks in orthants

Participant: Guy Fayolle.

We pursued works initiated these last years in several directions.

6.21.1. Explicit criterion for the finiteness of the group in the quarter plane

In the book [3], original methods were proposed to determine the invariant measure of random walks in the quarter plane with small jumps, the general solution being obtained via reduction to boundary value problems. Among other things, an important quantity, the so-called *group of the walk*, allows to deduce theoretical features about the nature of the solutions. In particular, when the *order* of the group is finite, necessary and sufficient conditions have been given in [3] for the solution to be rational or algebraic. When the underlying algebraic curve is of genus 1, we propose, in collaboration with R. Iasnogorodski (St-Petersbourg, Russia), a concrete criterion ensuring the finiteness of the group. It turns out that this criterion is always tantamount to the cancellation of a single constant, which can be expressed as the determinant of a matrix of order 3 or 4, and depends in a polynomial way on the coefficients of the walk [55].

6.21.2. About a possible analytic approach for walks with arbitrary big jumps in $\mathbb{Z}_+^2$

The article [21], achieved in collaboration with K. Raschel (CNRS and University F. Rabelais, Tours) considers random walks with arbitrary big jumps. For that class of models, we announce a possible extension of the analytic approach proposed in [3], initially valid for walks with small steps in the quarter plane. New technical challenges arise, most of them being tackled in the framework of generalized boundary value problems on compact Riemann surfaces.

6.21.3. Correction of papers

Guy Fayolle found important errors in several articles dealing with models involving random walks in $\mathbb{Z}_+^2$. This is the object of the letter to the editors [19]. The concerned authors have provided new correct versions of their studies.

6.21.4. Communication networks with harvesting energy supply

In collaboration with S. Foss (Heriot-Watt University, Edinburgh), we started to analyze stability and performance of a number of models of parallel queues with multiple access and individual energy supplies. Energy limitation in general decreases the stability region, but also may increase it for specific parameter regions. The most difficult and intriguing cases arise when the input rates of requests and of energy items are close. Preliminary models of physical interest involve random walks in $\mathbb{Z}_+^4$.

6.22. Global optimization for online resource allocation

Participant: Jean-Marc Lasgouttes.

As part of the Mobility 2.0 FP7 project, we have considered the possibility to allocate charging stations to Full Electric Vehicle (FEV) users in a way that, instead of merely minimizing their travel time, tries to improve the travel time for the whole community.

The aim of the global optimization algorithm is to pursue the minimization of the mean squared travel time encountered by each user. Our setting can be seen as a resource allocation problem, known as the "Transportation Problem" in Operational Research literature. It is solvable using several algorithms, among which the simplex algorithm. Unfortunately, these algorithms are not usable here for two reasons:

- The allocation of slots to the users is done online, when the user does a request. It is not possible to wait until all the users are known before doing the allocation;
- The complexity of these algorithms is very high, especially since, due to the effect of range limitations, each request has different characteristics.

We therefore present here a simplified approach, which should be fast enough to scale for large systems. The principle of the algorithm is to penalize the cost for the user with an approximation of the extra cost incurred to future users who compete for the same resource (a charging or parking slot). Since the implications can be intricate, we only consider a first order effect.

Our work in the Mobility 2.0 project has been briefly described in [37].

RMOD Project-Team

6. New Results

6.1. Highlights of the Year

- Pharo 3.0 has been released in April 2014.
- Moose 5.0 has been released in December 2014.
- The book Deep into Pharo has been released publicly http://www.deepintopharo.com.
- RMOD entered in a sponsoring agreement with LAM Research, Inc.

6.2. Tools for understanding applications

Remodularization Analysis Using Semantic Clustering. We report an experience on using and adapting Semantic Clustering to evaluate software remodularizations. Semantic Clustering is an approach that relies on information retrieval and clustering techniques to extract sets of similar classes in a system, according to their vocabularies. We adapted Semantic Clustering to support remodularization analysis. We evaluate our adaptation using six real-world remodularizations of four software systems. We report that Semantic Clustering and conceptual metrics can be used to express and explain the intention of the architects when performing common modularization operators, such as module decomposition. [37]

Towards a new package dependency model. Smalltalk originally did not have a package manager. Each Smalltalk implementation defined its own with more or less functionalities. Since 2010, Monticello/Metacello[Hen09] one package manager is available for open-source Smalltalks. It allows one to load source code packages with their dependencies. This package manager does not have all features we can find in well-known package managers like those used for the Linux operating system. We identify the missing features and propose a solution to reach a full-featured package manager. A part of this solution is to represent packages and dependencies as first-class objects, leading to the definition of a new dependency model. [32]

A Domain Specific Aspect Language for IDE Events. Integrated development environments (IDEs) have become the primary way to develop software. Besides just using the built-in features, it becomes more and more important to be able to extend the IDE with new features and extensions. Plugin architectures exist, but they show weaknesses related to unanticipated extensions and event handling. We argue that a more general solution for extending IDEs is needed. We present and discuss a solution, motivated by a set of concrete examples: a domain specific aspect language for IDE events. In it, join points are events of interest that may trigger the advice in which the behavior of the IDE extension is called. We show how this allows for the development of IDE plugins and demonstrate the advantages over traditional publish/subscribe systems. [21]

AspectMaps: Extending Moose to visualize AOP software. When using aspect-oriented programming the application implicitly invokes the functionality contained in the aspects. Consequently program comprehension of such a software is more intricate. To alleviate this difficulty we developed the AspectMaps visualization and tool. AspectMaps extends the Moose program comprehension and reverse engineering platform with support for aspects, and is implemented using facilities provided by Moose. We present the AspectMaps tool, and show how it can be used by performing an exploration of a fairly large aspect-oriented application. We then show how we extended the FAMIX meta-model family that underpins Moose to also provide support for aspects. This extension is called ASPIX, and thanks to this enhancement Moose can now also treat aspect-oriented software. Finally, we report on our experiences using some of the tools in Moose; Mondrian to implement the visualization, and Glamour to build the user interface. We discuss how we were able to implement a sizable visualization tool using them and how we were able to deal with some of their limitations. [20]

6.3. Software Quality: Taming Software Evolution

APIEvolutionMiner: Keeping API Evolution under Control. During software evolution, source code is constantly refactored. In real-world migrations, many methods in the newer version are not present in the old version (e.g., 60% of the methods in Eclipse 2.0 were not in version 1.0). This requires changes to be consistently applied to reflect the new API and avoid further maintenance problems. We propose a tool to extract rules by monitoring API changes applied in source code during system evolution. In this process, changes are mined at revision level in code history. Our tool focuses on mining invocation changes to keep track of how they are evolving. We also provide three case studies in order to evaluate the tool. [34]

Towards an Automation of the Mutation Analysis Dedicated to Model Transformation. A major benefit of Model Driven Engineering (MDE) relies on the automatic generation of artefacts from high-level models through intermediary levels using model transformations. In such a process, the input must be well-designed and the model transformations should be trustworthy. Due to the specificities of models and transformations, classical software test techniques have to be adapted. Among these techniques, mutation analysis has been ported and a set of mutation operators has been defined. However, mutation analysis currently requires a considerable manual work and suffers from the test data set improvement activity. This activity is seen by testers as a difficult and time-consuming job, and reduces the benefits of the mutation analysis. This work addresses the test data set improvement activity. Model transformation traceability in conjunction with a model of mutation operators, and a dedicated algorithm allow to automatically or semi-automatically produce test models that detect new faults. The proposed approach is validated and illustrated in a case study written in Kermeta. [17]

Predicting software defects with causality tests. We propose a defect prediction approach centered on more robust evidences towards causality between source code metrics (as predictors) and the occurrence of defects. More specifically, we rely on the Granger causality test to evaluate whether past variations in source code metrics values can be used to forecast changes in time series of defects. Our approach triggers alarms when changes made to the source code of a target system have a high chance of producing defects. We evaluated our approach in several life stages of four Java-based systems. We reached an average precision greater than 50% in three out of the four systems we evaluated. Moreover, by comparing our approach with baselines that are not based on causality tests, it achieved a better precision. [19]

6.4. Software Quality: History and Changes

Tracking dependencies between code changes: An incremental approach. Merging a change often leads to the question of knowing what are the dependencies to other changes that should be merged too to obtain a working system. This question also arises with code history trackers – Code history trackers are tools that react to what the developer do by creating first-class objects that represent the change made to the system. We evaluate the capacity of different code history trackers to represent, also as first-class objects, the dependencies between those changes. We also present a representation for dependencies that works with the event model of Epicea, a fine-grained and incremental code history tracker. [32]

Mining Architectural Violations from Version History. Software architecture conformance is a key software quality control activity that aims to reveal the progressive gap normally observed between concrete and planned software architectures. However, formally specifying an architecture can be difficult, as it must be done by an expert of the system having a high level understanding of it. We present a lightweighted approach for architecture conformance based on a combination of static and historical source code analysis. The proposed approach relies on four heuristics for detecting absences (something expected was not found) and divergences (something prohibited was found) in source code based architectures. We also present an architecture conformance process based on the proposed approach. We followed this process to evaluate the architecture of two industrial-strength information systems, achieving an overall precision of 62.7% and 53.8%. We also evaluated our approach in an open-source information retrieval library, achieving an overall precision of 59.2%. We envision that aheuristic-based approach for architecture conformance can be used to rapidly raise architectural warnings, without deeply involving experts in the process. [22]

6.5. Reconciling Dynamic Languages and Isolation

Delegation Proxies: The Power of Propagation. Scoping behavioral variations to dynamic extents is useful to support non-functional requirements that otherwise result in cross-cutting code. Unfortunately, such variations are difficult to achieve with traditional reflection or aspects. We show that with a modification of dynamic proxies, called delegation proxies, it becomes possible to reflectively implement variations that propagate to all objects accessed in the dynamic extent of a message send. We demonstrate our approach with examples of variations scoped to dynamic extents that help simplify code related to safety, reliability, and monitoring. [38]

Reifying the Reflectogram. Reflective facilities in OO languages are used both for implementing language extensions (such as AOP frameworks) and for supporting new programming tools and methodologies (such as object-centric debugging and message-based profiling). Yet controlling the run-time behavior of these reflective facilities introduces several challenges, such as computational overhead, the possibility of meta-recursion and an unclean separation of concerns between base and meta-level. We present five dimensions of meta-level control from related literature that try to remedy these problems. These dimensions are namely: temporal and spatial control, placement control, level control and identity control. We argue that the reification of the descriptive notion of the reflectogram, can unify the control of meta-level execution in all these five dimensions. We present a model for the reification of the reflectogram and validate our approach through a prototype implementation in the Pharo programming environment. Finally we detail a case study on run-time tracing illustrating our approach. [35]

Bootstrapping Reflective Systems: The Case of Pharo. Bootstrapping is a technique commonly known by its usage in language definition by the introduction of a compiler written in the same language it compiles. This process is important to understand and modify the definition of a given language using the same language, taking benefit of the abstractions and expression power it provides. A bootstrap, then, supports the evolution of a language. However, the infrastructure of reflective systems like Smalltalk includes, in addition to a compiler, an environment with several self-references. A reflective system bootstrap should consider all its infrastructural components. We propose a definition of bootstrap for object-oriented reflective systems, we describe the architecture and components it should contain and we analyze the challenges it has to overcome. Finally, we present a reference bootstrap process for a reflective system and Hazelnut, its implementation for bootstrapping the Pharo Smalltalk-inspired system. [26]

6.6. Dynamic Languages: Virtual Machines

Benzo: Reflective Glue for Low-level Programming. The goal of high-level low-level programming is to bring the abstraction capabilities of high-level languages to the system programming domain, such as virtual machines (VMs) and language runtimes. However, existing solutions are bound to compilation time and expose limited possibilities to be changed at runtime and from language-side. They do not fit well with fully reflective languages and environments. We propose Benzo1, a lightweight framework for high-level low-level programming that allows developers to generate and execute at runtime low-level code. It promotes the implementation, and dynamic modification, of system components with high-level language tools outperforming existing dynamic solutions. Since Benzo is a general framework we choose three applications that cover an important range of the spectrum of system programming for validating the infrastructure: a Foreign Function Interface (FFI), primitives instrumentation and a just-in-time bytecode compiler (JIT). With Benzo we show that these typical VM-level components are feasible as reflective language-side implementations. Due to its unique combination of high-level reflection and low-level programming, Benzo shows better performance for these three applications than the comparable high-level implementations. [30]

A bytecode set for adaptive optimizations. The Cog virtual machine features a bytecode interpreter and a baseline Just-in-time compiler. To reach the performance level of industrial quality virtual machines such as Java HotSpot, it needs to employ an adaptive inlining com-piler, a tool that on the fly aggressively optimizes frequently executed portions of code. We decided to implement such a tool as a bytecode to bytecode optimizer, implemented above the virtual machine, where it can be written and developed in Smalltalk. The optimizer we plan needs to extend the operations encoded in the bytecode set and its quality heavily depends on the bytecode set quality. The current bytecode set understood by the virtual machine is old and lacks any room to add new operations. We decided to implement a new bytecode set, which includes additional bytecodes that allow the Just-in-time compiler to generate less generic, and hence simpler and faster code sequences for frequently executed primitives. The new bytecode set includes traps for validating speculative inlining de-cisions and is extensible without compromising optimization opportunities. In addition, we took advantage of this work to solve limitations of the current bytecode set such as the maximum number of instance variable per class, or number of literals per method. We plan to have it in production in the Cog virtual machine and its Pharo, Squeak and Newspeak clients in the coming year. [43]

6.7. Traits

Trait-oriented Programming in Java 8 Java 8 was released recently. Along with lambda expressions, a new language construct is introduced: default methods in interfaces. The intent of this feature is to allow interfaces to be extended over time preserving backward compatibility. We show a possible, different use of interfaces with default methods: we introduce a trait-oriented programming style based on an interface-as-trait idea, with the aim of improving code modularity. Starting from the most common operators on traits, we introduce some programming patterns mimicking such operators and discuss this approach. [29]

6.8. Tailoring Applications

In the context of the PhD of G. Polito, we developed Tornado, a way to generate specialized and minimal runtime. Using a run-fail-grow approach, which tries to execute an expression in an empty world, and on failure copies the missing program elements from a mother environment to the currently empty world, we could grew 11k full reflective application adding two numbers or 18k for the 100 factorial expression. We also usd this approach to generate specialized webserver in around 500kb. These results show that we can generate hyperspecialized kernels.

ROMA Team

6. New Results

6.1. Highlights of the Year

Yves Robert was awarded the 2014 IEEE Technical Committee on Scalable Computing (TCSC) Award for Excellence.

In October 2014, CERFACS, ENS Lyon, INPT, Inria and University of Bordeaux launched a consortium around the software package MUMPS (see http://mumps-consortium.org).

6.2. Cost-Optimal Execution of Boolean DNF Trees with Shared Streams

Several applications process queries expressed as trees of Boolean operators applied to predicates on sensor data streams, e.g., mobile apps and automotive apps. Sensor data must be retrieved from the sensors, which incurs a cost, e.g., an energy expense that depletes the battery of a mobile device, a bandwidth usage. The objective is to determine the order in which predicates should be evaluated so as to shortcut part of the query evaluation and minimize the expected cost. This problem has been studied assuming that each data stream occurs at a single predicate. In this work [17], [27] we study the case in which a data stream occurs in multiple predicates, either when each predicate references a single stream or when a predicate can reference multiple streams. In the single-stream case we give an optimal algorithm for a single-level tree and show that the problem is NP-complete for DNF trees. For DNF trees we show that there exists an optimal predicate evaluation order that is depth-first, which provides a basis for designing a range of heuristics. In the multi-stream case we show that the problem is NP-complete even for single-level trees. As in the single stream case, for DNF trees we show that there exists a depth-first leaf evaluation order that is optimal and we design efficient heuristics.

6.3. Efficient checkpoint/verification patterns for silent error detection

Errors have become a critical problem for high performance computing. Checkpointing protocols are often used for error recovery after fail-stop failures. However, silent errors cannot be ignored, and their peculiarity is that such errors are identified only when the corrupted data is activated. To cope with silent errors, we need a verification mechanism to check whether the application state is correct. Checkpoints should be supplemented with verifications to detect silent errors. When a verification is successful, only the last checkpoint needs to be kept in memory because it is known to be correct. In this work (RR UT-EECS-14-729), we analytically determine the best balance of verifications and checkpoints so as to optimize platform throughput. We introduce a balanced algorithm using a pattern with p checkpoints and q verifications, which regularly interleaves both checkpoints and verifications across same-size computational chunks. We show how to compute the waste of an arbitrary pattern, and we prove that the balanced algorithm is optimal when the platform MTBF (Mean Time Between Failures) is large in front of the other parameters (checkpointing, verification and recovery costs). We conduct several simulations to show the gain achieved by this balanced algorithm for well-chosen values of p and q, compared to the base algorithm that always perform a verification just before taking a checkpoint ($p = q = 1$), and we exhibit gains of up to 19%.

6.4. Assessing general-purpose algorithms to cope with fail-stop and silent errors

In this work (RR-Inria-8599), we combine the traditional checkpointing and rollback recovery strategies with verification mechanisms to address both fail-stop and silent errors. The objective is to minimize either makespan or energy consumption. While DVFS is a popular approach for reducing the energy consumption, using lower speeds/voltages can increase the number of errors, thereby complicating the problem. We consider an application workflow whose dependence graph is a chain of tasks, and we study three execution scenarios: (i) a single speed is used during the whole execution; (ii) a second, possibly higher speed is used for any potential re-execution; (iii) different pairs of speeds can be used throughout the execution. For each scenario, we determine the optimal checkpointing and verification locations (and the optimal speeds for the third scenario) to minimize either objective. The different execution scenarios are then assessed and compared through an extensive set of experiments.

6.5. Scheduling the I/O of HPC applications under congestion

A significant percentage of the computing capacity of large-scale platforms is wasted due to interferences incurred by multiple applications that access a shared parallel file system concurrently. One solution to handling I/O bursts in large-scale HPC systems is to absorb them at an intermediate storage layer consisting of burst buffers. However, our analysis of the Argonne's Mira system shows that burst buffers cannot prevent congestion at all times. As a consequence, I/O performance is dramatically degraded, showing in some cases a decrease in I/O throughput of 67%. In this work (RR-Inria-8519), we analyze the effects of interference on application I/O bandwidth, and propose several scheduling techniques to mitigate congestion. We show through extensive experiments that our global I/O scheduler is able to reduce the effects of congestion, even on systems where burst buffers are used, and can increase the overall system throughput up to 56%. We also show that it outperforms current Mira I/O schedulers.

6.6. Power-aware replica placement in tree networks with multiple servers per client

In this work (RR-Inria-8474), we revisit the well-studied problem of replica placement in tree networks. Rather than minimizing the number of servers needed to serve all client requests, we aim at minimizing the total power consumed by these servers. In addition, we use the most general (and powerful) server assignment policy, where the requests of a client can be served by multiple servers located in the (unique) path from this client to the root of the tree. We consider multi-modal servers that can operate at a set of discrete speeds, using the dynamic voltage and frequency scaling (DVFS) technique. The optimization problem is to determine an optimal location of the servers in the tree, as well as the speed at which each server is operated. A major result is the NP-completeness of this problem, to be contrasted with the minimization of the number of servers, which has polynomial complexity. Another important contribution is the formulation of a Mixed Integer Linear Program (MILP) for the problem, together with the design of several polynomial-time heuristics. We assess the efficiency of these heuristics by simulation. For mid-size instances (up to 30 nodes in the tree), we evaluate their absolute performance by comparison with the optimal solution (obtained via the MILP). The most efficient heuristics provide satisfactory results, within 20% of the optimal solution.

6.7. Parallel scheduling of task trees with limited memory

This work [28] investigates the execution of tree-shaped task graphs using multiple processors. Each edge of such a tree represents some large data. A task can only be executed if all input and output data fit into memory, and a data can only be removed from memory after the completion of the task that uses it as an input data. Such trees arise, for instance, in the multifrontal method of sparse matrix factorization. The peak memory needed for the processing of the entire tree depends on the execution order of the tasks. With one processor the objective of the tree traversal is to minimize the required memory. This problem was well studied and optimal polynomial algorithms were proposed.

Here, we extend the problem by considering multiple processors, which is of obvious interest in the application area of matrix factorization. With multiple processors comes the additional objective to minimize the time needed to traverse the tree, i.e., to minimize the makespan. Not surprisingly, this problem proves to be much harder than the sequential one. We study the computational complexity of this problem and provide inapproximability results even for unit weight trees. We design a series of practical heuristics achieving different trade-offs between the minimization of peak memory usage and makespan. Some of these heuristics are able to process a tree while keeping the memory usage under a given memory limit. The different heuristics are evaluated in an extensive experimental evaluation using realistic trees.

6.8. Scheduling Trees of Malleable Tasks for Sparse Linear Algebra

Scientific workloads are often described as directed acyclic task graphs. In this work [30], we focus on the multifrontal factorization of sparse matrices, whose task graph is structured as a tree of parallel tasks. Among the existing models for parallel tasks, the concept of *malleable* tasks is especially powerful as it allows each task to be processed on a time-varying number of processors. Following the model advocated by Prasanna and Musicus [62], [63] for matrix computations, we consider malleable tasks whose speedup is p^{α}, where p is the fractional share of processors on which a task executes, and α ($0 < \alpha \leq 1$) is a parameter which does not depend on the task. We first motivate the relevance of this model for our application with actual experiments on multicore platforms. Then, we study the optimal allocation proposed by Prasanna and Musicus for makespan minimization using optimal control theory. We largely simplify their proofs by resorting only to pure scheduling arguments. Building on the insight gained thanks to these new proofs, we extend the study to distributed multicore platforms. There, a task cannot be distributed among several distributed nodes. In such a distributed setting (homogeneous or heterogeneous), we prove the NP-completeness of the corresponding scheduling problem, and propose some approximation algorithms. We finally assess the relevance of our approach by simulations on realistic trees. We show that the average performance gain of our allocations with respect to existing solutions (that are thus unaware of the actual speedup functions) is up to 16% for $\alpha = 0.9$ (the value observed in the real experiments).

6.9. Non-clairvoyant reduction algorithms for heterogeneous platforms

In this work [6], we have revisited the classical problem of the reduction collective operation in a heterogeneous environment. We have discussed and evaluated four algorithms that are non-clairvoyant, i.e., they do not know in advance the computation and communication costs. On the one hand, Binomial-stat and Fibonacci-stat are static algorithms that decide in advance which operations will be reduced, without adapting to the environment; they were originally defined for homogeneous settings. On the other hand, Tree-dyn and Non-Commut-Tree-dyn are fully dynamic algorithms, for commutative or non-commutative reductions. We have shown that these algorithms are approximation algorithms with constant or asymptotic ratios. We assessed the relative performance of all four non-clairvoyant algorithms with heterogeneous costs through a set of simulations. Our conclusions hold for a variety of distributions.

6.10. Memory-aware tree traversals with pre-assigned tasks

We have studied the complexity of traversing tree-shaped workflows whose tasks require large I/O files. We target a heterogeneous architecture with two resource types, each with a different memory, such as a multicore node equipped with a dedicated accelerator (FPGA or GPU). The tasks in the workflow are colored according to their type and can be processed if all there input and output files can be stored in the corresponding memory. The amount of used memory of each type at a given execution step strongly depends upon the ordering in which the tasks are executed, and upon when communications between both memories are scheduled. The objective is to determine an efficient traversal that minimizes the maximum amount of memory of each type needed to traverse the whole tree. In this study [11], we establish the complexity of this two-memory scheduling problem, and provide inapproximability results. In addition, we design several heuristics, based on both post-order and general traversals, and we evaluate them on a comprehensive set of tree graphs, including random trees as well as assembly trees arising in the context of sparse matrix factorizations.

6.11. Analysis of Dynamic Scheduling Strategies for Matrix Multiplication on Heterogeneous Platforms

The tremendous increase in the size and heterogeneity of supercomputers makes it very difficult to predict the performance of a scheduling algorithm. Therefore, dynamic solutions, where scheduling decisions are made at runtime have overpassed static allocation strategies. The simplicity and efficiency of dynamic schedulers such as Hadoop are a key of the success of the MapReduce framework. Dynamic schedulers such as StarPU, PaRSEC or StarSs are also developed for more constrained computations, e.g. task graphs coming from linear algebra. To make their decisions, these runtime systems make use of some static information, such as the distance of tasks to the critical path or the affinity between tasks and computing resources (CPU, GPU,. . .) and of dynamic information, such as where input data are actually located. In this study [16], we concentrate on two elementary linear algebra kernels, namely the outer product and the matrix multiplication. For each problem, we propose several dynamic strategies that can be used at runtime and we provide an analytic study of their theoretical performance. We prove that the theoretical analysis provides very good estimate of the amount of communications induced by a dynamic strategy and can be used in order to efficiently determine thresholds used in dynamic scheduler, thus enabling to choose among them for a given problem and architecture.

6.12. Determining the optimal redistribution

The classical redistribution problem aims at optimally scheduling communications when reshuffling from an initial data distribution to a target data distribution. This target data distribution is usually chosen to optimise some objective for the algorithmic kernel under study (good computational balance or low communication volume or cost), and therefore to provide high efficiency for that kernel. However, the choice of a distribution minimizing the target objective is not unique. This leads to generalizing the redistribution problem as follows: find a re-mapping of data items onto processors such that the data redistribution cost is minimal, and the operation remains as efficient. This work studies the complexity of this generalized problem. We compute optimal solutions and evaluate, through simulations, their gain over classical redistribution. We also show the NP-hardness of the problem to find the optimal data partition and processor permutation (defined by new subsets) that minimize the cost of redistribution followed by a simple computational kernel. Finally, experimental validation of the new redistribution algorithms are conducted on a multicore cluster, for both a 1D-stencil kernel and a more compute-intensive dense linear algebra routine.

6.13. On the hierarchically structured bin packing problem

We study the hierarchically structured bin packing problem [14]. In this problem, the items to be packed into bins are at the leaves of a tree. The objective of the packing is to minimize the total number of bins into which the descendants of an internal node are packed, summed over all internal nodes. We investigate an existing algorithm and make a correction to the analysis of its approximation ratio. Further results regarding the structure of an optimal solution and a strengthened inapproximability result are given.

6.14. Heuristics for the bipartite matching problem

We propose two heuristics for the bipartite matching problem that are amenable to shared-memory parallelization [18]. The first heuristic is very intriguing from parallelization perspective. It has no significant algorithmic synchronization overhead and no conflict resolution is needed across threads. We show that this heuristic has an approximation ratio of around 0.632. The second heuristic is designed to obtain a larger matching by employing the well-known Karp-Sipser heuristic on a judiciously chosen subgraph of the original graph. We show that the Karp-Sipser heuristic always finds a maximum cardinality matching in the chosen subgraph. Although the Karp-Sipser heuristic is hard to parallelize for general graphs, we exploit the structure of the selected subgraphs to propose a specialized implementation which demonstrates a very good scalability. Based on our experiments and theoretical evidence, we conjecture that this second heuristic obtains matchings with cardinality of at least 0.866 of the maximum cardinality. We discuss parallel implementations of the proposed heuristics on shared memory systems. Experimental results, for demonstrating speed-ups and verifying the theoretical results in practice, are provided.

6.15. Fill-in reduction in sparse matrix factorizations using hypergraphs

We discuss the use of hypergraph partitioning based methods in fill-reducing orderings of sparse matrices for Cholesky, LU and QR factorizations [33]. For the Cholesky factorization, we investigate a recent result on pattern-wise decomposition of sparse matrices, generalize the result, and develop algorithmic tools to obtain more effective ordering methods. The generalized results help us formulate the fill-reducing ordering problem for LU factorization as we do for the Cholesky case, without ever symmetrizing the given matrix A as $|A| + |A^T|$ or $|A^T||A|$. For the QR factorization, we adopt a recently proposed technique to use hypergraph models in a fairly standard manner. The method again does not form the possibly much denser matrix $|A^T||A|$. We also discuss alternatives for LU and QR factorization cases where the symmetrized matrix can be used. We provide comparisons with the most common alternatives in all three cases.

6.16. On partitioning two dimensional finite difference meshes for distributed memory parallel computers

We investigate the problem of partitioning finite difference meshes in two dimensions among the processors of a parallel computer [20]. The objective is to achieve a perfect load balance while minimizing the communication cost. There are well-known graph, hypergraph, and geometry-based partitioning algorithms for this problem. The known geometric algorithms have linear running time and obtain the best results for very special mesh sizes and processor numbers. We propose another geometric algorithm. The proposed algorithm is linear; is applicable to much more cases than some well-known alternatives; obtains better results than the graph partitioning algorithms; obtains better results than the hypergraph partitioning algorithms almost always. Our algorithm also obtains better results than a known asymptotically-optimal algorithm for some small number of processors. We also catalog related theoretical results.

6.17. A symmetry preserving algorithm for matrix scaling

We present an iterative algorithm which asymptotically scales the ∞-norm of each row and each column of a matrix to one [12]. This scaling algorithm preserves symmetry of the original matrix and shows fast linear convergence with an asymptotic rate of 1/2. We discuss extensions of the algorithm to the one-norm, and by inference to other norms. For the 1-norm case, we show again that convergence is linear, with the rate dependent on the spectrum of the scaled matrix. We demonstrate experimentally that the scaling algorithm improves the conditioning of the matrix and that it helps direct solvers by reducing the need for pivoting. In particular, for symmetric matrices the theoretical and experimental results highlight the potential of the proposed algorithm over existing alternatives.

6.18. Direct solvers for sparse linear systems

In the context of the MUMPS sparse direct solver (see Section 5.1), we worked in 2014 on: block-low-rank solvers and shared-memory parallelism [4], [13], hybrid (shared-distributed) parallelism and efficient collective communications in asynchronous environments [2], and scheduling strategies to decrease the memory-usage of multifrontal solvers. Quite significant performance gains have been obtained on up to 2000 cores of a Bullx DLC system (CALMIP mesocentre), some of the corresponding developments will be made available in the next release of our solver. We also worked on setting up a consortium of industrial users to fund engineers working on MUMPS (see Section 7.1). These activities were done in collaboration with INP Toulouse and with CERFACS, CNRS, ENS Lyon, Univ. Bordeaux, EDF, LSTC (Livermore, California) and EMGS (Norway).

RUNTIME Team

6. New Results

6.1. Highlights of the Year

- This year we started very large collaborations with the BULL/Atos company. WE started one European project, one PIA french project and one PhD thesis. The amount of Person Year funded with this project exceed 10. The research we will do with Bull covers resource management, process placement, platform modeling, application modeling, affinity abstraction.
- The StarPU software is used by CEA for automatically distributing linear algebra on their cluster of 144 hybrid nodes.

6.2. Task scheduling over heterogeneous architectures

We continued our work on extending STARPU to master exploitation of Heterogeneous Platforms through dynamic task scheduling, with a now-imminent release of StarPU 1.2.

We have improved the simulation support with SIMGRID, to augment the accuracy of the simulated execution according to the hardware capabilities [30].

We have collaborated with various research projects to leverage the potential of STARPU. We have improved the support for the PASTIX and QR-MUMPS sparse matrix solvers, thus obtaining competitive performance on CPUs and on CPUs+GPUs [25]. We have improved the MPI communication engine of STARPU to get better performance with the EADS hmatrix solver.

We have obtained very good performance and scalability with a Cholesky factorization distributed over a cluster of 144 heterogeneous nodes hosted at CEA.

We have studied the theoretical performance bound that can be achieved for the Cholesky factorization, reproduced the performance of a theorically optimal scheduled, shown that the classical HEFT heuristic is far from it, that more application-specific heuristics allow to get performance closer to the peak, and that the peak is not reachable with simple heuristics, because it requires non-trivial task order inversions.

In relationship with the ADT K'Star effort of building the KLANG-OMP OpenMP compiler and putting together the KASTORS benchmark suite, StarPU has been extended to provide an OpenMP-enabled runtime support for KLANG-OMP. In particular, the StarPU OpenMP Runtime Support implements *preemptible* tasks required for OpenMP, using the concept of continuations, while maintaining interoperability with StarPU regular, non-blocking tasks, and while preserving the heterogeneous, performance model-based scheduling capabilities of StarPU.

The KLANG-OMP C/C++ OpenMP compiler co-developed with Inria Team MOAIS enables plain OpenMP applications to run un-modified on top of the StarPU runtime system, thus significantly increasing the performance portability potential of StarPU.

6.3. Modeling hierarchical platform memory performance with microbenchmarks

Bertrand PUTIGNY developed a new memory performance model based on micro-benchmarks during his PhD. He transforms parallel codes such as OpenMP into memory access skeleton before predicting memory buffer states in caches and using benchmarks outputs to predict the runtime. This model successfully predict the performance behavior of several memory-bound kernels [26].

We also used this model to study the impact of memory caches on the performance on intra-node MPI communication [27].

6.4. Static modeling of clusters of multicore and heterogeneous nodes

We improved the hwloc software to better manage clusters of nodes. This first includes the management of HPC node I/O devices by providing easy ways to retrieve the locality of GPUs and network interfaces. A scalable global view of clusters can be built by factorizing the common topology information that is usually shared by many similar nodes [20]. Finally the topology of the network assembling all these nodes can be exposed in a generic technology-independent manner using the new netloc tool [21] that is now part of hwloc.

6.5. Multithreaded communications

We have proposed a full rewrite of the PIOMAN software, to make it rely on system threads rather than on the now obsolete MARCEL thread scheduler. It makes it more portable, composable with any runtime system used for multithreading, and more scalable. We have shown [19][18] that it features good properties with regard to asynchronous communication progression and multithreaded communications in applications.

6.6. Toplogy-aware load balancing in Charm++

Charm++ implements a fine-grained paradigm based on migratable computing objects. This programming model is designed to run large-scale experiments and provide a dynamic load balancing system to optimize it. Our previous Charm++ load balancer designed for communication-bound applications was improved to scale on large platforms. More precisely, we worked on the network awareness of this algorithm by using LibTopoMap. Our topology-aware load balancing algorithm was also restructured to be parallel and distributed. These enhancements were validated on the Blue Waters supercomputer at Urbana-Champaign, IL. Finally, We have begun to carry out experiments on real application modeling seismic wave propagation.

6.7. Topology-aware ressource allocation

On the one hand SLURM already provides topology aware placement techniques to promote the choice of group of nodes that are placed on the same network level, connected under the same network switch or even placed close to each other so as to avoid long distance communications. On the other hand users can map tasks in a parallel application to the physical processors on the chosen nodes, based on the communication topology.

Our goal is to take in account, in SLURM, placement process, hardware topology, and application communication pattern too. We have implemented a new selection option for the cons_res plugin in SLURM 2.6.5. In this case the usually best fit algorithm used to choose nodes is replaced by Treematch, an algorithm to find the best placement among the free nodes list in light of a given application communication matrix. Tests and evaluation of this feature are in progress.

6.8. Scheduling of dynamic streaming applications on hybrid embedded MPSoCs

The work on the dataflow scheduler has continued so as to improve it: it is now simper and more efficient. Moreover, an H.264 video decoder implementation from STMicroelectronics has been ported onto the developed execution model to conduct more significant experiments. This application exhibits a higher level of complexity and variability, which is the reason why it is well suited for assessing the scheduler's reactivity. Furthermore, an important groundwork has been carried out to enable software support for parts of the application, which enlarges considerably the design space and allow to benefit from better flexibility. In parallel, some earlier work on list scheduling under memory constraints has been extended and published in an international journal [11].

6.9. Performance model for multithreaded applications on multi-core processors

Concerning data locality, researches have shown a tradeof in groupement strategy for process mapping. We have to deal with balanced improvement of several aspects such as threads synchronizations or resource exploitation. Weighting those criterias can only be achieved according to a certain knowledge of both the application and the machine.

Thus, we are working on modeling threads affinity and weights on machines topology to improve a placement method based on the TreeMatch algorithm using new metrics. Several experiences have lead us to the conclusion that it is very hard to identify the key hints and to understand application needs.

Consequently, we are developping a visual tool which displays hardware counters aggregated and mapped on the system topology to identify dynamically those hardware narrows during execution, and understand processes placement effects on them. We hope to achieve a better comprehension of process placement consequences on resources usage by applications.

SAGE Project-Team

6. New Results

6.1. Highlights of the Year

Lionel Lenôtre and his co-authors revisited in a very efficient way the Hastings-Metropolis Algorithm on Markov Chains for Small-Probability Estimation.

6.2. Numerical algorithms

6.2.1. Hybrid algebraic sparse linear solvers

Participants: Jocelyne Erhel, David Imberti.

Grants and projects: EXA2CT 7.3.1 , C2S@EXA 7.2.3

Publications: [17]

Abstract: Sparse linear systems arise in computational science and engineering. The goal is to reduce the memory requirements and the computational cost, by means of high performance computing algorithms. Krylov methods combined with Domain Decomposition are very efficient. Numerical results show the benefits of our methodology.

6.2.2. GMRES and Polynomial Equivalence

Participant: David Imberti.

Grants and projects: EXA2CT 7.3.1 , C2S@EXA 7.2.3

Publications: in preparation.

Abstract: We have established a theoretical link between GMRES and the much simpler problem of polynomial evaluation along with some algebraic structures to describe the most important elements of the GMRES algorithm. We use these structures to show the connection between sequential GMRES and Horner's Rule, s-step GMRES and Dorn's rule, and predict future possible GMRES-like algorithms.

6.2.3. Variables s-step GMRES

Participant: David Imberti.

Grants and projects: EXA2CT 7.3.1 , C2S@EXA 7.2.3

Publications: in preparation.

Abstract: We introduce a new variation on s-step GMRES in order to improve its stability, reduce the number of iterations necessary to ensure convergence, and thereby improve parallel performance. In doing so, we develop a new block variant that allows us to express the stability difficulties in s-step GMRES more fully. We use the algebraic structures previous established via the polynomial equivalence to support an intuitive choice for the variation in the s-step procedure, and reinforce its utility in some communication cost estimates.

6.2.4. FGMRES dynamics

Participant: David Imberti.

Grants and projects: EXA2CT 7.3.1 , C2S@EXA 7.2.3

Publications: in preparation.

Abstract: The FGMRES algorithm has met with varying success and we detail theoretical relationships between FGMRES and GMRES including a geometric mean conjecture. Further, we build on the current literature regarding GMRES convergence with an analysis of the dynamical properties of FGMRES.

6.2.5. RPM Coupling Factors

Participant: David Imberti.

Grants and projects: EXA2CT 7.3.1 , C2S@EXA 7.2.3

Publications: in preparation.

Abstract: We have improved the Recursive Projection Method (RPM) with a subspace version that effectively utilizes parallelism. Furthermore, we include a discussion, numerical experiments, and suggestions for the heretofor neglected coupling factor in RPM, and how they influence convergence of the algorithm.

6.2.6. Hastings-Metropolis Algorithm on Markov Chains for Small-Probability Estimation

Participant: Lionel Lenôtre.

Grants: H2MNO4 7.2.1

Publications: [13]

Abstract: Shielding studies in neutron transport, with Monte Carlo codes, yield challenging problems of small-probability estimation. The particularity of these studies is that the small probability to estimate is formulated in terms of the distribution of a Markov chain, instead of that of a random vector in more classical cases. Thus, it is not straightforward to adapt classical statistical methods, for estimating small probabilities involving random vectors, to these neutron-transport problems. A recent interacting-particle method for small-probability estimation, relying on the Hastings-Metropolis algorithm, is presented. It is shown how to adapt the Hastings-Metropolis algorithm when dealing with Markov chains. A convergence result is also shown. Then, the practical implementation of the resulting method for small-probability estimation is treated in details, for a Monte Carlo shielding study. Finally, it is shown, for this study, that the proposed interacting-particle method considerably outperforms a simple Monte Carlo method, when the probability to estimate is small.

6.2.7. A Strategy for the Parallel Implementations of Stochastic Lagrangian Methods

Participant: Lionel Lenôtre.

Grants: H2MNO4 7.2.1

Software: PALMTREE 5.3.1

Publications: [34]

Abstract: We present some investigations on the parallelization of a stochastic Lagrangian simulation. For the self sufficiency of this work, we start by recalling the stochastic methods used to solve Parabolic Partial Differential Equations with a few physical remarks. Then, we exhibit different object-oriented ideas for such methods. In order to clearly illustrate these ideas, we give an overview of the library PALMTREE that we developed. After these considerations, we discuss the importance of the management of random numbers and argue for the choice of a particular strategy. To support our point, we show some numerical experiments of this approach, and display a speedup curve of PALMTREE. Then, we discuss the problem in managing the parallelization scheme. Finally, we analyze the parallelization of hybrid simulation for a system of Partial Differential Equations. We use some works done in hydrogeology to demonstrate the power of such a concept to avoid numerical diffusion in the solution of Fokker-Planck Equations and investigate the problem of parallelizing scheme under the constraint entailed by domain decomposition. We conclude with a presentation of the latest design that was created for PALMTREE and give a sketch of the possible work to get a powerful parallelized scheme.

6.3. Numerical models and simulations applied to physics

6.3.1. Small scale modeling of porous media

Participants: Édouard Canot, Salwa Mansour.

Grants: ARPHYMAT 7.4.3 , 7.4.4

Software: GLiMuH 5.4.3

Publications: [22]

Abstract: This study is devoted to the heat transfer between two spherical grains separated by a small gap; dry air is located around the grains and a liquid water meniscus is supposed to be present between them. This problem can be seen as a micro-scale cell of an assembly of solid grains, for which we are looking for the effective thermal conductivity. For a fixed contact angle and according to the volume of the liquid meniscus, two different shapes are possible for the meniscus, giving a "contacting" state (when the liquid makes a true bridge between the two spheres) and a "non-contacting" one (when the liquid is split in two different drops, separated by a thin air layer); the transition between these two states occurs at different times when increasing or decreasing the liquid volume, thus leading to a hysteresis behavior when computing the thermal flux across the domain.

6.3.2. Heat and mass transfer modeling in porous media

Participants: Édouard Canot, Salwa Mansour.

Grants: HYDRINV 7.4.5

Software: HeMaTiS (5.4.1)

Abstract: The physical model of the HeMaTiS code (5.4.1) has been recently improved by adding the diffusion process of dry air through the water steam which is created by the evaporation of the water inside the porous medium. In this fashion, not only can the heating stage of the surface of the soil be simulated but also the cooling stage. The application concerns the study of archaeological fires which were used many times a day; the possibility of alternation of heating and cooling may lead to a better interpretation of residual marks left in the ground. Work is in progress to validate the numerical results.

6.3.3. Inverse problem for determining the thermo-physical properties of a porous media

Participants: Édouard Canot, Salwa Mansour.

Grants: HYDRINV 7.4.5

Software: TPIP (5.4.2)

Publications: [23]

Abstract: This study concerns the inverse problem which consists of the estimation of thermophysical properties of the soil knowing the temperature at few selected points of the domain. In order to solve this inverse problem, we used the least square criterion where we try to minimize the error function between real measures and simulated ones. The coupled system composed of the energy equation together with the three sensitivity boundary initial problems resulting from differentiating the basic energy equation with respect to the soil properties must be solved. To overcome the stiffness of our problem (due to the use of Apparent Heat Capacity method), the high nonlinearity of the coupled system and the problem of large residuals we used the Damped Gauss Newton and Levenberg-Marquardt methods. Moreover, we emphasized on the importance of the choice of ΔT (temperature range over which the phase change occurs) where for a certain initial guess the inverse problem fails to converge. We overcome this problem by chaining the inverse problems using different values of ΔT and parameters' set.

6.3.4. Geodesy

Participant: Bernard Philippe.

Grants: LIRIMA-EPIC 7.4.2 .

Publications: [12].

Abstract: We solve a geodetic inverse problem for the determination of a distribution of point masses (characterized by their intensities and positions), such that the potential generated by them best approximates a given potential field.

6.4. Models and simulations for flow and transport in porous media

6.4.1. Simulating Diffusion Processes in Discontinuous Media: Benchmark Tests

Participant: Géraldine Pichot.

Grants: H2MN04 7.2.1

Software: SBM 5.2.2

Publications: [33]

Abstract: We present several benchmark tests for Monte Carlo methods for simulating diffusion in one-dimensional discontinuous media, such as the ones arising the geophysics and many other domains. These benchmarks tests are developed according to their physical, statistical, analytic and numerical relevance. We then perform a systematic study on four numerical methods.

6.4.2. Uncertainty Quantification and High Performance Computing for flow and transport in porous media

Participants: Jean-Raynald de Dreuzy, Jocelyne Erhel.

Grants: HYDRINV 7.4.5 , H2MN04 7.2.1

Publications: [18]

Abstract: Stochastic models use random fields to represent heterogeneous porous media. Quantities of interest such as macro dispersion are then analyzed from a statistical point of view. In order to get asymptotic values, large scale simulations must be performed, using High Performance Computing. Non-intrusive methods are well-suited for two-level parallelism. Indeed, for each simulation, parallelism is based on domain decomposition for generating the random input and the flow matrix, parallel linear solvers and parallel particle tracker. Also, several simulations, corresponding to randomly drawn samples, can run in parallel. The balance between these two levels depends on the resources available. The software PARADIS implements flow and transport with random data and computation of macro dispersion. Simulations run on supercomputers with large 3D domains.

6.4.3. Computation of macro spreading in 3D porous media with uncertain data

Participants: Jean-Raynald de Dreuzy, Jocelyne Erhel, Mestapha Oumouni.

Grants: HYDRINV 7.4.5 , H2MN04 7.2.1

Publications: [15]

Abstract: We consider an heterogeneous porous media where the conductivity is described by probability laws. Thus the velocity, which is solution of the flow equation, is also a random field, taken as input in the transport equation of a solute. The objective is to get statistics about the spreading and the macro dispersion of the solute. We use a mixed finite element method to compute the velocity and a lagrangian method to compute the spreading. Uncertainty is dealt with a classical Monte-Carlo method, which is well-suited for high heterogeneities and small correlation lengths. We give an explicit formulation of the macro dispersion and a priori error estimates. Numerical experiments with large 3D domains are done with the software PARADIS of the platform H2OLab.

6.4.4. A combined collocation and Monte-Carlo method for advection-diffusion equation of a solute in random porous media

Participants: Jocelyne Erhel, Mestapha Oumouni.

Grants: HYDRINV 7.4.5 , H2MN04 7.2.1

Publications: [14]

Abstract: In this work, we present a numerical analysis of a method which combines a deterministic and a probabilistic approaches to quantify the migration of a contaminant, under the presence of uncertainty on the permeability of the porous medium. More precisely, we consider the flow equation in a random porous medium coupled with the advection-diffusion equation. Quantities of interest are the mean spread and the mean dispersion of the solute. The means are approximated by a quadrature rule, based on a sparse grid defined by a truncated Karhunen-Loève expansion and a stochastic collocation method. For each grid point, the flow model is solved with a mixed finite element method in the physical space and the advection-diffusion equation is solved with a probabilistic Lagrangian method. The spread and the dispersion are expressed as functions of a stochastic process. A priori error estimates are established on the mean of the spread and the dispersion.

6.4.5. An adaptive sparse grid method for elliptic PDEs with stochastic coefficients

Participants: Jocelyne Erhel, Mestapha Oumouni.

Grants: HYDRINV 7.4.5 , H2MN04 7.2.1

Publications: [31].

Abstract: The stochastic collocation method based on the anisotropic sparse grid has become a significant tool to solve partial differential equations with stochastic inputs. The aim is to seek a vector of weights and a convenient level of interpolation for the method. The classical approach uses an a posteriori approach on the solution, which causes an additional prohibitive cost.

In this work, we discuss an adaptive approach of this method to calculate the statistics of the solution. It is based on an adaptive approximation of the *inverse* diffusion parameter. We construct an efficient error indicator which is an upper bound of the error on the solution. In the case of unbounded variables, we use an appropriate error estimation to compute suitable weights for the method. Numerical examples are presented to confirm the efficiency of the approach, and to show that the cost is considerably reduced without loss of accuracy.

6.4.6. Numerical analysis of stochastic advection-diffusion equation via Karhunen-Loève expansion

Participants: Jocelyne Erhel, Mestapha Oumouni.

Grants: HYDRINV 7.4.5 , H2MN04 7.2.1

Publications: [32] , [25]

Abstract: In this work, we present a convergence analysis of a probabilistic approach to quantify the migration of a contaminant, under the presence of uncertainty on the permeability of the porous medium. More precisely, we consider the flow problem in a random porous medium coupled with the advection-diffusion equation and we are interested in the approximation of the mean spread and the dispersion of the solute. The conductivity field is represented by a Karhunen-Loève (K-L) decomposition of its logarithm. The flow model is solved using a mixed finite element method in the physical space. The advection-diffusion equation is computed thanks to a probabilistic Lagrangian method, where the concentration of the solute is the density function of a stochastic process. This process is solution of a stochastic differential equation (SDE), which is discretized using an Euler scheme. Then, the mean of the spread and dispersion are expressed as functions of the approximate stochastic process. A priori error estimates are established on the mean of the spread and of the dispersion. Numerical examples show the effectiveness of this approach.

6.4.7. About a generation of a log-normal correlated field

Participants: Jocelyne Erhel, Mestapha Oumouni.

Grants: HYDRINV 7.4.5 , H2MN04 7.2.1

Software: GENFIELD 5.2.3

Publications: in preparation

Abstract: Uncertainty quantification often requires the generation of large realizations of stationary Gaussian random field over a regular grid.

This paper compares and analyzes the commonalities between the methods and approaches for simulating stationary Gaussian random field. The continuous spectral method is the classical approach which discretizes its spectral density to construct an approximation of the field. When the spectral density and the covariance functions decrease rapidly to zero at infinity, we prove that the spectral method is computationally attractive.

We compare also the classical methods used to simulate the field defined by its covariance function, namely the Discrete Spectral method, the Circulant Embedding approach, and the Discrete Karhunen-Loève approximation. We have found that under some assumptions on the covariance all these latter methods give the same simulations of a stationary Gaussian field on a regular grid, which are very efficient with the Fast Fourier Transform algorithm.

6.4.8. A global model for reactive transport

Participants: Édouard Canot, Jocelyne Erhel.

Grants: H2MN04 7.2.1 , MOMAS 7.2.5 , C2SEXA 7.2.3

Software: GRT3D 5.2.1

Thesis: [11]

Publications: [19], [16]

Abstract: In some scientific applications, such as groundwater studies, several processes are represented by coupled models. For example, numerical simulations are essential for studying the fate of contaminants in aquifers, for risk assessment and resources management. Chemical reactions must be coupled with advection and dispersion when modeling the contamination of aquifers. This coupled model combines partial differential equations with algebraic equations, in a so-called PDAE system, which is nonlinear. A classical approach is to follow a method of lines, where space is first discretized, leading to a semi-discrete differential algebraic system (DAE). Several methods have been designed for solving this system of PDAE.

In this study, we propose a global method which uses a DAE solver, where time is discretized by an implicit scheme. Then, each time step involves a nonlinear system of equations, solved by a modified Newton method. Thanks to the DAE solver, the time step is adaptively chosen in order to ensure accuracy and convergence. Moreover, the Jacobian in the nonlinear iterations is freezed as long as Newton converges fast enough, saving a lot of CPU time.

However, the size of the nonlinear system is quite large, because it involves both the differential and the algebraic variables. We show how to eliminate the differential variables, in order to reduce the size. This is equivalent to a so-called Direct Substitution Approach, but it keeps the nice features of DAE solvers.

Classicaly, the concentrations of chemical species are defined with their logarithms, assuming that they are strictly positive. This simplifies the computation of the mass action laws in the chemistry model and the computation of their derivatives. However, when a species does not exist, its concentration is replaced by a very small value and this may lead to an ill-conditioned Jacobian. We propose to use directly the concentrations, without logarithms, so that the Jacobian is then well-conditioned. Therefore, Newton method converges much faster without logarithms, allowing larger time steps and saving many computations.

We illustrate our method with two test cases, provided by the french agency for nuclear waste (ANDRA) and by the group MOMAS. We can compare our results with either analytical or other numerical solutions and show that our method is quite accurate. We also show that reducing the number of unknowns is very efficient and that dealing without logarithms reduces drastically the CPU time.

6.4.9. A chemistry model with precipitation-dissolution

Participant: Jocelyne Erhel.

Grants: H2MN04 7.2.1 , MOMAS 7.2.5

Internship: Tangi Migot (Master M2, INSA and University, Rouen)

Publications: [36]

Abstract: In this study, we focus on precipitation and dissolution chemical reactions, because they induce numerical difficulties.

We consider a set of solute species and minerals, with precipitation occuring when a saturation threshold is reached. A challenge is to detect which minerals are dissolved and which minerals are precipitated. This depends on the total quantities of chemical species. We propose an analytical approach to build a phase diagram, which provides the interfaces between the different possible cases. We illustrate our method with three examples arising from brine media and acid mine drainage.

6.4.10. Coupled models for salted aquifers

Participants: Édouard Canot, Jocelyne Erhel.

Grants: H2MN04 7.2.1 , MOMAS 7.2.5 , HYDRINV 7.4.5

Software: GEODENS and SELSAUM (from Tunis)

Internship: Marwen ben Refifa (Ph-D, ENIT, Tunis)

Publications: in preparation

Abstract: We study gravity driven problems in salted aquifers, when many species are present together with high concentrations. In this framework, we couple flow, transport and chemistry by using a fixed point approach. We interfaced two codes developed in Tunis: GEODENS for density driven flow and transport, and SELSAUM for geochemistry. This latter provides also the density of salted water.

6.5. Models and simulations for flow in porous fractured media

6.5.1. A Graphical User Interface for simulating flow and transport in fractured-porous media

Participants: Jean-Raynald de Dreuzy, Jocelyne Erhel, Géraldine Pichot.

Grants: H2MNO4 7.2.1

Platform: H2OLab

Publications: [21]

Abstract: The platform H2OLab can be used with a Graphical User Interface, called H2OGuilde, which is developed using a Qt framework. Launchers correspond to a main program and to a hydrogeological application. These launchers call modules or libraries, implementing discretization schemes, solving algorithms, parallel communications, etc. The interface is generic for all the launchers. It is composed of three main tabs corresponding to the three steps of a simulation: entering input data, running computations, analyzing output data. Input parameters are classified in several categories, corresponding to the physical model and the numerical algorithms chosen. Output parameters are of three types, scalar, vector and matrix. Currently, visualization is done outside of the interface.

6.5.2. Meshing Strategies and the Impact of Finite Element Quality on the Velocity Field in Fractured Media

Participants: Jean-Raynald de Dreuzy, Jocelyne Erhel, Géraldine Pichot.

Grants: FRACINI 7.1.1

Platform: H2OLab

Publications: [20]

Abstract: For solving flow within a network of fractures, Mixed Hybrid Finite Element (MHFE) method is a method of choice as it yields a symmetric positive definite linear system. However, a drawback is its sensitivity to bad aspect ratio elements. For poor-quality triangles, elementary matrices are ill-conditioned and inconsistent velocity vectors are obtained by inverting these local matrices. In our presentation, we will present different strategies for a better reconstruction of the velocity field.

SCALE Team

6. New Results

6.1. Programming Languages for Distributed Systems

One of the objectives of the Scale team is to design programming models easing the development and safe execution of distributed systems. This section describes our results in this direction.

6.1.1. Multi-active Objects

Participants: Ludovic Henrio, Fabrice Huet, Justine Rochas, Vincenzo Mastandrea.

The active object programming model is particularly adapted to easily program distributed objects: it separates objects into several *activities*, each manipulated by a single thread, preventing data races. However, this programming model has its limitations in terms of expressiveness – risk of deadlocks – and of efficiency on multicore machines. We proposed to extend active objects with *local multi-threading*. We rely on declarative *annotations* for expressing potential concurrency between requests, allowing easy and high-level expression of concurrency. This year we realized the following:

- We published the extension of multi-active objects to support scheduling and thread limitation [12].
- We developed a compiler from ABS language into ProActive multi-active objects. This translation can be generalised to many other active object languages. This work has been published as a research report [25], and is under submission to a conference.
- We started to work on static detection of deadlocks for multi-active object. This is the work of Vincenzo Mastandrea who is starting a Labex PhD in collaboration with the FOCUS EPI (Univ of Bologna).
- Extensive use of multiactive objects in our CAN P2P network and implementation of usecases [2].
- We formalised in Isabelle/HOL a first version of the semantics of multiactive objects. This work was done in collaboration with Florian Kammuller

We plan to continue to improve the model, especially about compile-time checking of annotations and about fault tolerance of multiactive objects.

6.1.2. Autonomic Monitoring and Management of Components

Participants: Françoise Baude, Ludovic Henrio.

We have completed the design of a framework for autonomic monitoring and management of component-based applications. We have provided an implementation using GCM/ProActive taking advantage of the possibility of adding components in the membrane. The framework for autonomic computing allows the designer to describe in a separate way each phase of the MAPE autonomic control loop (Monitoring, Analysis, Planning, and Execution), and to plug them or unplug them dynamically.

- This year, we published a journal paper summarising our approach in the GCM/ProActive framework and our contribution on componentised membranes for autonomic computing [3].
- We also improved, in the context of the SCADA associate team and during the internship of Matias Ibañez, the support for autonomic components, providing all the architecture and API so that the programmer of autonomic aspect can do them in a DSL reconfiguration language, called GCMScript. This was implemented and experimented, a publication is under submission on this work.

6.1.3. Algorithmic skeletons

Participant: Ludovic Henrio.

In the context of the SCADA associated team, we worked on the algorithmic skeleton programming model. The structured parallelism approach (skeletons) takes advantage of common patterns used in parallel and distributed applications. The skeleton paradigm separates concerns: the distribution aspect can be considered separately from the functional aspect of an application. In the previous year we designed the possibility for a skeleton to output events, which increases the control and monitoring capabilities. This year we published our previous results in [14] and realised additional steps:

- Study of different ways to predict the execution time for a skeleton, inspired from simple statistic functions. This improvement together with the distributed execution of skeletons should allow us to publish a journal paper on this subject in 2015

6.1.4. Optimization of data transfer in event-based programming models

Participants: Iyad Alshabani, Françoise Baude, Laurent Pellegrino.

In [6], we extended a previous work with conceptual and experimental performance evaluations. This previous and collaborative work [1] developed an innovative approach of "lazy copy and transfer" of the data parts of event objects exchanged by peers in the context of event-driven architecture applications.

While event notifications are routed in a conventional manner through an event service, data parts of the events are directly and transparently transferred from publishers to subscribers. The theoretical analysis shows that we can reduce the average event delivery time by half, compared to a conventional approach requiring the full mediation of the event service. The experimental analysis confirms that the proposed approach outperforms the conventional one (both for throughput and delivery time) even though the middleware overhead, introduced by the specific adopted model, slightly reduces the expected benefits.

6.1.5. Behavioural Semantics

Participants: Ludovic Henrio, Eric Madelaine, Min Zhang.

We have studied Parameterised Networks of Automata (pNets) from a theoretical perspective. We started with some 'pragmatic' expressiveness of the pNets formalism, showing how to express a wide range of classical constructs of (value-passing) process calculi, but also complex interaction patterns used in modern distributed systems. Our framework can model full systems, using (closed) hierarchies of pNets; we can also build (open) pNet systems expressing composition operators. Concerning more fundamental aspects, we defined a strong bisimulation theory specifically for the pNet model, proved its properties, and illustrated it on some examples. One of the original aspects of the approach is to relate the compositional nature of pNets with the notion of bisimulation; this was exemplified by studying the properties of a flattening operator for pNets. This work has been accepted for publication at PDP'2015 ([24]).

6.1.6. A Time-sensitive Heterogeneous Behavioural Model

Participants: Eric Madelaine, Yanwen Chen.

This work concludes the PhD research of Yanwen Chen, targetting a timed-sensitive extension of the pNets model with logical clocks inherited from the CCSL language. The main results of this year are: 1) a new notion of Time Specification (TS), used to handle the abstract properties of each level of processes in a pNet structure, 2) algorithms to compute such TSs for basic parameterized and timed processes, and from composition of timed-pNets, 3) conditions for checking the compatibility of composition, 4) a use-case from the area of intelligent transportation systems, illustrating the whole chain of modeling, upto a symbolic simulation of the full composed system, with the TimeSquare tool. This work was published as [4], [23], and in the PhD thesis of Y. Chen, defended on 2014, Nov. 30th.

6.1.7. Structure and structural correctness for GCM components

Participants: Ludovic Henrio, Oleksandra Kulankhina, Eric Madelaine.

We have defined a set of rules characterizing the well-formed composition of components in order to guarantee their safe deployment and execution. This work focuses on the structural aspects of component composition; it puts together most of the concepts common to many component models, but never formalized as a whole. Our formalization characterizes correct component architectures made of functional and non-functional aspects, both structured as component assemblies. So-called 'Interceptor chains' can be used for a safe and controlled interaction between the two aspects. Our well-formed components guarantee a set of properties ensuring that the deployed component system has a correct architecture and can run safely. Those definitions constitute the formal basis for VerCors tool. This work was done in the context of O. Kulankhina phd research, and in collaboration with Dongqian Liu (ECNU Shanghai), as part of the Associated Team DAESD.

6.2. Run-time/middle-ware level

6.2.1. Scalable and robust Middleware for distributed event based computing

Participants: Françoise Baude, Fabrice Huet, Laurent Pellegrino, Maeva Antoine.

In the context of the FP7 STREP PLAY and French SocEDA ANR research projects terminated late 2013, we initiated and pursued the design and development of the Event Cloud. This has been the core content of Laurent Pellegrino PhD thesis [2], and the corresponding software deposit at the APP for this middleware.

As a distributed system, this middleware can suffer from failures. To resist to such situations, we have added a capability of checkpointing. In [18] we present how to design an adaptation of the famous Chandy and Lamport algorithm for distributed snapshot taking, to the case of the Event Cloud. Indeed, as the Event Cloud peers are multi-active objects, we need to take care when and how to serve the chekpointing request and so, when to apply the Chandy Lamport protocol operations. Consequently, we have make sure that the obtained distributed snapshot is indeed consistent. As publication of events are triggered from the outside of the Event Cloud, we however are not able to recover them from the last saved snapshot in case of peer crash and subsequent whole Event Cloud recovery. However, we ensure any event injected through a peer, before this peer was participating in the last global checkpoint taking is safely part of it.

As a distributed system handling huge amount of information, this middleware can suffer from data imbalances. In [22], [8], we have reviewed the litterature of structured peer to peer systems regarding the way they handle load imbalance. We have generalized those popular approaches by proposing a core API that we have proved to be indeed also applicable to the Event Cloud middleware way of implementing a load balancing policy.

Storing highly skewed data in a distributed system has become a very frequent issue, in particular with the emergence of semantic web and big data. This often leads to biased data dissemination among nodes. Addressing load imbalance is necessary, especially to minimize response time and avoid workload being handled by only one or few nodes. We have proposed a protocol which allows a peer to change its hash function at runtime, without a priori knowledge regarding data distribution. This provides a simple but efficient adaptive load balancing mechanism. Moreover, we have shown that a structured overlay can still be consistent event when all peers do not apply the same hash function on data [7].

6.2.2. Virtual Machines Placement Algorithms

Participants: Fabien Hermenier, Vincent Kherbache.

In [21], [19], we present BtrPlace as an application of the dynamic bin packing problem with a focus on its dynamic and heterogeneous nature. We advocate flexibility to answer these issues and present the theoretical aspects of BtrPlace and its modeling using Constraint Programming. In [5] we rely on BtrPlace to achieve energy efficiency. To maintain an energy footprint as low as possible, data centres manage their VMs according to conventional and established rules. Each data centre is however made unique due to its hardware and workload specificities. This prevents the *ad-hoc* design of current VM schedulers from taking these particularities into account to provide additional energy savings. In this paper, we present Plug4Green, an application that relies on BtrPlace to customize an energy-aware VM scheduler. This flexibility is validated through the implementation of 23 SLA constraints and 2 objectives aiming at reducing either the power consumption or the greenhouse gas emissions. On a heterogeneous test bed, Plug4Green specialization to fit the hardware and the workload specificities allowed to reduce the energy consumption and the gas emission by up to 33% and 34%, respectively. Finally, simulations showed that Plug4Green is capable of computing an improved placement for 7,500 VMs running on 1,500 servers within a minute.

Finally, we started to investigate on easing the jobs of data centre operators using BtrPlace. For example, server maintenance is a common but still critical operation. A prerequisite is indeed to relocate elsewhere the VMs running on the production servers to prepare them for the maintenance. When the maintenance focuses several servers, this may lead to a costly relocation of several VMs so the migration plan must be chose wisely. This however implies to master numerous human, technical, and economical aspects that play a role in the design of a quality migration plan. In[13], we study migration plans that can be decided by an operator to prepare for an hardware upgrade or a server refresh on multiple servers. We exhibit performance bottleneck and pitfalls that reduce the plan efficiency. We then discuss and validate possible improvements deduced from the knowledge of the environment peculiarities.

6.3. Application level

6.3.1. Simulation Software Architecture

Participant: Olivier Dalle.

In general purpose software engineering (as opposed to simulation software engineering), the motivations for reuse have long been advocated and demonstrated: lower risks of defects, collective support of potentially larger user community, lower development costs, and so on. In simulation software architectures, we can also cite business-specific motivations, such as providing a better reproducibility of simulation experiments, or avoiding a complex validation process. In [20], we show that although it is rarely discussed, reuse is a problem that may be considered in two opposite directions: reusing and being reused.

6.3.2. DEVS-based Modeling & Simulation

Participants: Olivier Dalle, Damian Vicino.

DEVS is a formalism for the specification of discrete-event simulation models, proposed by Zeigler in the 70's, that is still the subject of many research in the simulation community. Surprisingly, the problem of representing the time in this formalism has always been somehow neglected, and most DEVS simulators keep using Floating Point numbers for their arithmetics on time values, which leads to a range of systematic errors, including severe ones such as breaking the causal relations in the model. In [16] we propose a new data type for discretized time representation in DEVS, based on rational numbers. Indeed, we show that rational numbers offer good stability properties for the arithmetics used in DEVS, with a limited impact on the simulation execution performance.

6.3.3. GPU-based High Performance Cloud Computing

Participants: Michael Benguigui, Françoise Baude, Fabrice Huet.

To address HPC, GPU devices are now considered as unavoidable cheap, energy efficient, and very efficient alternative computing units. Our long term goal is to devise some generic solutions in order to incorporate GPU-specific code whenever relevant into a parallel and distributed computation.

As a challenging example, we have pursued our work on pricing American multi-dimensional (so very computation intensive) options in finance. From our previous work that achieved pricing a 40-assets based american option within 8 hours of computation on a single GPU, the work in [9] allows us to reach approximatively one hour of computation time. For this, we run using active objects coupled with OpenCL codes, on 18 GPU nodes acquired from the Grid'5000 platform (the maximum amount of available GPU on Grid'5000 that we could book at once).

Moreover, the balancing of work is taking in consideration the heterogeneous nature of the involved GPUs, and is capable to harness the computing power of multi-core CPUs that also support running OpenCL codes. This parallel and distributed pricing approach is also extended in the forthcoming PhD thesis of Michael Benguigui: it successfully tackles the Value At Risk computation of a portofolio composed of such complex financial products.

6.3.4. Simulation of Software-Defined Networks

Participants: Olivier Dalle, Damian Vicino.

Software Defined Networks (SDN) is a new technology that has gained a lot of attention recently. It introduces programmatic ways to reorganize the network logical topology. To achieve this, the network interacts with a set of controllers, that can dynamically update the configuration of the network routing equipments based on the received events. As often with new network technologies, discrete-event simulation proves to be an invaluable tool for understanding and analzing the performance and behavior of the new systems. In [17], we use such smulations for evaluating the impact of Software-Defined Networks' Reactive Routing on BitTorrent performance. Indeed, BitTorrent uses choking algorithms that continuously open and close connections to different peers. Software Defined Networks implementing Reactive Routing may be negatively affecting the performances of the system under specific conditions because of it lack of knowledge of BitTorrent strategies.

SECRET Project-Team

6. New Results

6.1. Highlights of the Year

- Rafael Misoczki's PhD thesis on code-based cryptography (defended in November 2013) has been awarded by the Brazilian Society of Computer Science as the best thesis in computer security.
- *Security analysis of some primitives for authentication and authenticated encryption:* authentication is a major functionality in the vast majority of applications. It is usually implemented by a MAC (message authentication code). The main constructions for MAC are based on hash functions, and include the wide-spread HMAC construction. Gaëtan Leurent, together with Itai Dinur, has presented a new generic attack against HMAC when the underlying hash function follows the Haifa construction. This result points out that the hash function in HMAC has to be chosen very carefully and that some of the main families of hash functions may introduce unexpected weaknesses in the associated MAC. Also, the project-team is involved in a national cryptanalytic effort funded by the ANR which aims at evaluating the security of the recently proposed authenticated encryption schemes.
- *Parallel Repetition of Entangled Games:* In a two-player free game G, two cooperating but non communicating players receive inputs taken from two independent probability distributions. Each of them produces an output and they win the game if they satisfy some predicate on their inputs/outputs. The classical (resp. entangled) value of G is the maximum winning probability when the players are allowed to share classical random bits (resp. a quantum state) prior to receiving their inputs. The n-fold parallel repetition of G consists of n instances of G where the parties receive all the inputs at the same time, produce all the outputs at the same time and must win every instance of G. This work by André Chailloux in collaboration with Giannicola Scarpa establishes that the entangled value of the parallel repetition of G decreases exponentially with n, thereby generalizing to the quantum setting Raz's celebrated parallel repetition theorem which is concerned with the classical value of the game. The main tool for proving this result is the introduction of a new information-theoretic quantity: the superposed information cost.

6.2. Symmetric cryptosystems

Participants: Anne Canteaut, Pascale Charpin, Virginie Lallemand, Gaëtan Leurent, María Naya Plasencia, Joëlle Roué, Valentin Suder.

From outside, it might appear that symmetric techniques become obsolete after the invention of public-key cryptography in the mid 1970's. However, they are still widely used because they are the only ones that can achieve some major features like high-speed or low-cost encryption, fast authentication, and efficient hashing. Today, we find symmetric algorithms in GSM mobile phones, in credit cards, in WLAN connections. Symmetric cryptology is a very active research area which is stimulated by a pressing industrial demand for low-cost implementations (in terms of power consumption, gate complexity...). These extremely restricted implementation requirements are crucial when designing secure symmetric primitives and they might be at the origin of some weaknesses. Actually, these constraints seem quite incompatible with the rather complex mathematical tools needed for constructing a provably secure system.

The specificity of our research work is that it considers all aspects in the field, from the practical ones (new attacks, concrete specifications of new systems) to the most theoretical ones (study of the algebraic structure of underlying mathematical objects, definition of optimal objects). But, our purpose is to study these aspects not separately but as several sides of the same domain. Our approach mainly relies on the idea that, in order to guarantee a provable resistance to the known attacks and to achieve extremely good performance, a symmetric cipher must use very particular building blocks, whose algebraic structures may introduce unintended weaknesses. Our research work captures this conflict for all families of symmetric ciphers. It includes new attacks and the search for new building blocks which ensure both a high resistance to the known attacks and a low implementation cost. This work, which combines cryptanalysis and the theoretical study of discrete mathematical objects, is essential to progress in the formal analysis of the security of symmetric systems.

In this context, the very important challenges are the designs of low-cost ciphers and of authenticated encryption schemes. Most teams in the research community are actually working on the design and on the analysis (cryptanalysis and optimization of the performance) of such primitives.

6.2.1. Block ciphers

Even if the security of the current block cipher standard, AES, is not threatened when it is used in a classical context, there is still a need for the design of improved attacks, and for the determination of design criteria which guarantee that the existing attacks do not apply. This notably requires a deep understanding of all previously proposed attacks. Moreover, there is a high demand from the industry of lightweight block ciphers for some constrained environments. Several such algorithms have been proposed in the last few years and their security should be carefully analyzed. Most of our work in this area is related to an ANR Project named BLOC. Our recent results then mainly concern either the analysis and design of lightweight block ciphers, or the in-depth study of the security of the block cipher standard AES.

Recent results:

- Cryptanalysis of several recently proposed lightweight block ciphers. This includes an attack against the full cipher KLEIN-64 [60], an attack against 8 rounds (out of 12) of PRINCE [48], [77], and an attack against Zorro and its variants [74].
- Formalization and generic improvements of impossible differential cryptanalysis: this type of attacks, even if extensively used, remains not fully understood, and it appears that there are numerous applications where mistakes have been discovered or where the attacks lack optimality. Our work then provides a general framework for impossible differential cryptanalysis including a generic complexity analysis of the optimal attack. Using these advances, we have also presented the best known impossible differential attacks against several ciphers including CLEFIA-128, Camellia, LBlock and Simon [46], [76], [75].
- Design of a new family of block ciphers achieving very good software performance, especially on 8-bit microcontrollers. A nice feature of these ciphers is that they offer an optimal resistance against side-channel attacks in the sense that the cost of Boolean masking is minimized [58].
- Design and study of a new construction for low-latency block ciphers, named *reflection ciphers*, which generalizes the so-called α-reflection property exploited in PRINCE. This construction aims at reducing the implementation overhead of decryption on top of encryption [24]
- Proposal of a new family of distinguishers against AES-based permutations, named *limited-birthday distinguishers*; these distinguishers exploit some some improved rebound techniques. They have been successfully applied to various AES-based primitives including AES, ECHO, Grøstl, LED, PHOTON and Whirlpool [18].
- Analysis of the differential and linear properties of the AES Superbox [65].

6.2.2. Authenticated encryption

A limitation of all classical block ciphers is that they aim at protecting confidentiality only, while most applications need both encryption and authentication. These two functionalities are provided by using a block cipher like the AES together with an appropriate mode of operation. However, it appears that the most widely-used mode of operation for authenticated encryption, AES-GCM, is not very efficient for high-speed networks. Also, the security of the GCM mode completely collapses when an IV is reused. These severe drawbacks have then motivated an international competition named CAESAR, partly supported by the NIST, which has been recently launched in order to define some new authenticated encryption schemes [0]. Our work related to this competition is then two-fold: G. Leurent has participated to the design of a CAESAR candidate named SCREAM. Also, the project-team is involved in a national cryptanalytic effort led by the BRUTUS project funded by the ANR which aims at evaluating the security of all CAESAR candidates.

Recent results:

- Submission of a proposal to the CAESAR competition [88], [67].
- Cryptanalysis of three CAESAR candidates: Wheesht [64], π-cipher [90], LAC [69].

6.2.3. Hash functions and MACS

The international research effort related to the selection of the new hash function standard SHA-3 has led to many important results and to a better understanding of the security offered by hash functions. However, hash functions are used in a huge number of applications with different security requirements, and also form the building-blocks of some other primitives, like MACs. In this context, we have investigated the security of some of these constructions, in order to determine whether some particular constructions for hash functions may affect the security of the associated MACs.

Recent results:

- Improved generic attacks against hash-based MAC, including HMAC, when the hash function follows the Haifa construction [55], [33];
- Attack against Streebog, the new Russian hash function standard: we show that the specific instantiation of the Haifa construction used in Streebog makes it weak against second-preimage attacks [59].

6.2.4. Cryptographic properties and construction of appropriate building blocks

The construction of building blocks which guarantee a high resistance against the known attacks is a major topic within our project-team, for stream ciphers, block ciphers and hash functions. The use of such optimal objects actually leads to some mathematical structures which may be at the origin of new attacks. This work involves fundamental aspects related to discrete mathematics, cryptanalysis and implementation aspects. Actually, characterizing the structures of the building blocks which are optimal regarding to some attacks is very important for finding appropriate constructions and also for determining whether the underlying structure induces some weaknesses or not.

For these reasons, we have investigated several families of filtering functions and of S-boxes which are well-suited for their cryptographic properties or for their implementation characteristics. For instance, bent functions, which are the Boolean functions which achieve the highest possible nonlinearity, have been extensively studied in order to provide some elements for a classification, or to adapt these functions to practical cryptographic constructions. We have also been interested in functions with a low differential uniformity (*e.g.*, APN functions), which are the S-boxes ensuring an (almost) optimal resistance to differential cryptanalysis.

[0] http://competitions.cr.yp.to/caesar.html

Recent results:

- Study of the algebraic properties (e.g. the algebraic degree) of the inverses of APN power permutations [19].
- Study of the cryptographic properties, including the degree, the differential uniformity and the size of the image set of permutations of the form $x \mapsto x^s + \gamma \mathsf{Tr}(x^t)$ over a finite field of characteristic two [15]. Since these functions are obtained by slightly modifying a power function, they share similar interesting implementation properties but do not present the weaknesses coming from their structure. In particular, an infinite family of permutations of this form with differential uniformity 4 has been exhibited.
- Definition of an extended criterion for estimating the resistance of a block cipher to differential attacks. Most notably, this new criterion points out the fact that affinely equivalent Sboxes may not provide the same security level regarding differential and linear cryptanalysis. This work emphasizes the role played by the affine permutation of the set of 8-bit words which follows the inverse function in the AES [65].

6.2.5. *Symmetric primitives based on lattices*

Lattice-based cryptography is an alternative to number-theoretic constructions for public-key cryptography. Lattice-based constructions enjoy a worst-case security reduction to hard lattice problems, and the area is very active, with many new designs offering attractive features.

Recently, this approach has also been used to build symmetric cryptosystems based on lattice problems. While those systems are less efficient than traditional symmetric systems, they are still reasonably efficient, and their security can be related to hard computational problems rather than being only heuristic. In addition, the underlying mathematical structure can offer extra properties such as parallelizability or easy protection against side-channel attacks.

Recent results:

- Design of a family of pseudo-random functions named SPRING which aims to combine the guarantees of security reductions with good performance [44]; implementation of SPRING on FPGA and protection of this hardware implementation against side-channel attacks [47].
- Implementation and side-channel evaluation of the Lapin authentication protocol, based on the LPN problem [57].

6.3. Code-based cryptography

Participants: Julia Chaulet, Adrien Hauteville, Grégory Landais, Nicolas Sendrier, Jean-Pierre Tillich.

Most popular public-key cryptographic schemes rely either on the factorization problem (RSA, Rabin), or on the discrete logarithm problem (Diffie-Hellman, El Gamal, DSA). These systems have evolved and today instead of the classical groups ($\mathbf{Z}/n\mathbf{Z}$) we may use groups on elliptic curves. They allow a shorter block and key size for the same level of security. An intensive effort of the research community has been and is still being conducted to investigate the main aspects of these systems: implementation, theoretical and practical security. It must be noted that these systems all rely on algorithmic number theory. As they are used in most, if not all, applications of public-key cryptography today (and it will probably remain so in the near future), cryptographic applications are thus vulnerable to a single breakthrough in algorithmics or in hardware (a quantum computer can break all those schemes).

Diversity is a way to dilute that risk, and it is the duty of the cryptographic research community to prepare and propose alternatives to the number-theoretic-based systems. The most serious tracks today are lattice-based cryptography (NTRU,...), multivariate cryptography (HFE,...) and code-based cryptography (McEliece encryption scheme,...). All these alternatives are referred to as *post-quantum cryptosystems*, since they rely on difficult algorithmic problems which would not be solved by the coming-up of the quantum computer.

The code-based primitives have been investigated in details within the project-team. The first cryptosystem based on error-correcting codes was a public-key encryption scheme proposed by McEliece in 1978; a dual variant was proposed in 1986 by Niederreiter. We proposed the first (and only) digital signature scheme in 2001. Those systems enjoy very interesting features (fast encryption/decryption, short signature, good security reduction) but also have their drawbacks (large public key, encryption overhead, expensive signature generation). Some of the main issues in this field are

- security analysis, implementation and practicality of existing solutions,
- reducing the key size, *e.g.*, by using rank metric instead of Hamming metric, or by using particular families of codes,
- addressing new functionalities, like hashing or symmetric encryption.

Recent results:

- Cryptanalysis of McEliece system based on Wild Goppa codes from a quadratic finite field extension. This polynomial-time structural attack relies on some filtration of nested subcodes which will reveal the secret algebraic description of the underlying secret code [16], [17].
- Structural cryptanalysis of some variants of McEliece scheme based on alternant codes which have a quasi-cyclic or quasi-dyadic generator matrix [86].
- Cryptanalysis of a variant of the McEliece cryptosystem based on Reed-Solomon codes [16].
- Design of a new variant of McEliece using quasi-cyclic Moderate Density Parity Check (MDPC) codes [39].

6.4. Reverse-engineering of communication systems

Participants: Marion Bellard, Nicolas Sendrier, Jean-Pierre Tillich, Audrey Tixier.

To assess the quality of a cryptographic algorithm, it is usually assumed that its specifications are public, as, in accordance with Kerckhoffs principle [0], it would be dangerous to rely, even partially, on the fact that the adversary does not know those specifications. However, this fundamental rule does not mean that the specifications are known to the attacker. In practice, before mounting a cryptanalysis, it is necessary to strip off the data. This reverse-engineering process is often subtle, even when the data formatting is not concealed on purpose. A typical case is interception; some raw data, not necessarily encrypted, are observed out of a noisy channel. To access the information, the whole communication system has first to be disassembled and every constituent reconstructed. Our activity within this domain, whose first aim is to establish the scientific and technical foundations of a discipline which does not exist yet at an academic level, has been supported by some industrial contracts driven by the Ministry of Defense.

Recent results:

- Reconstruction of the constellation labelling (i.e. used in the modulator of a communication system) in the presence of errors and when the underlying code is convolutional [10].
- Reconstruction of a convolutional code. This reconstruction technique is based on a new method for detecting whether a given binary sequence is a noisy convolutional codeword obtained from an unknown convolutional code [45].
- Reconstruction of the interleaver of a turbo-code from the knowledge of several noisy codewords [63].

6.5. Quantum information theory

Participants: André Chailloux, Anthony Leverrier, Denise Maurice, Jean-Pierre Tillich.

[0] Kerckhoffs stated that principle in a paper entitled *La Cryptographie militaire*, published in 1883.

The field of Quantum Information and Computation aims at exploiting the laws of quantum physics to manipulate information in radically novel ways. Two main applications come to mind: quantum computers, that offer the promise of solving some problems intractable with classical computers (for instance, factorization); and quantum cryptography, which provides new ways to exchange data in a provably secure fashion.

The main obstacle towards the development of quantum computing is decoherence, a consequence of the interaction of the computer with a noisy environment. We investigate approaches to quantum error-correction as a way to fight against this effect, and we study more particularly some families of quantum error-correcting codes which generalize the best classical codes available today.

Our research also covers quantum cryptography where we study the security of efficient protocols for key distribution or coin flipping, in collaboration with experimental groups. More generally, we investigate how quantum theory severely constrains the action of honest and malicious parties in cryptographic scenarios.

Finally, a promising approach to better understand the possibilities of quantum information consists in studying quantum correlations via the notion of nonlocal games, where different parties need to coordinate to answer some questions, but without communicating. The goal here is to analyze the optimal strategies and to quantify the quantum advantage, i.e. how much sharing an entangled quantum state helps compared to sharing classical randomness.

6.5.1. Quantum codes

Protecting quantum information from external noise is an issue of paramount importance for building a quantum computer. It also worthwhile to notice that all quantum error-correcting code schemes proposed up to now suffer from the very same problem that the first (classical) error-correcting codes had: there are constructions of good quantum codes, but for the best of them it is not known how to decode them in polynomial time. Our approach for overcoming this problem has been to study whether or not the family of turbo-codes and LDPC codes (and the associated iterative decoding algorithms) have a quantum counterpart.

Recent results:

- Construction of quantum LDPC codes with fixed non-zero rate and a minimum distance which grows proportionally to the square root of the block-length. This greatly improves the previously best known construction whose minimum distance was logarithmic in the block-length [23].
- Design of a decoding algorithm for the family of quantum codes due to Calderbank, Shor and Steane [84].
- Study of quantum error correcting codes with an iterative decoding algorithm [12].
- Error analysis for Boson Sampling, a simplified model for quantum computation [91].

6.5.2. Quantum cryptography

A recent approach to cryptography takes into account that all interactions occur in a physical world described by the laws of quantum physics. These laws put severe constraints on what an adversary can achieve, and allow for instance to design provably secure key distribution protocols. We study such protocols as well as more general cryptographic primitives such as coin flipping with security properties based on quantum theory.

Recent results:

- Composable security proof for a continuous-variable quantum key distribution protocol with coherent states [92], [71], [70].
- Proof of existence of quantum weak coin flipping with arbitrarily small bias [80].
- Experimental implementation of quantum coin flipping [20].
- Study of connections between quantum encodings, non-locality and quantum cryptography [22].

6.5.3. Quantum correlations and nonlocality

Since the seminal work from Bell in the 60's, it has been known that classical correlations obtained via shared randomness cannot reproduce all the correlations obtained by measuring entangled quantum systems. This impossibility is for instance witnessed by the violation of a Bell inequality and is known under the name of "Quantum Nonlocality". In addition to its numerous applications for quantum cryptography, the study of quantum nonlocality and quantum games has become a central topic in quantum information theory, with the hope of bringing new insights to our understanding of quantum theory.

Recent results:

- Proof of parallel repetition of entangled games with exponential decay [52],[82],[32].
- Development of a general framework for the study of quantum correlations with combinatorial tools [35].
- New bounds on the quantum value of nonlocal games with graph-theoretical arguments [51].
- Optimal bounds for parity-oblivious random access codes [50].
- Study of Local Orthogonality, a physical principle upper bounding quantum correlations [21].
- Considerations on the notion of dimension of physical systems and its implications for information processing [14].

SELECT Project-Team

6. New Results

6.1. Model selection in Regression and Classification

Participants: Gilles Celeux, Serge Cohen, Clément Levrard, Erwan Le Pennec, Pascal Massart, Nelo Molter Magalhaes, Lucie Montuelle.

Unsupervised segmentation is an issue similar to unsupervised classification with an added spatial aspect. Functional data is acquired on points in a spatial domain and the goal is to segment the domain in homogeneous domain. The range of applications includes hyperspectral images in conservation sciences, fMRi data and all spatialized functional data. Erwan Le Pennec and Lucie Montuelle are focusing on the questions of the way to handle the spatial component from both the theoretical and the practical point of views. They study in particular the choice of the number of clusters. Furthermore, as functional data require heavy computation, they are required to propose numerically efficient algorithms. With Serge Cohen and an X intern some progress have been made on the use of logistic weights in the hyperspectral setting.

Lucie Montuelle has studied a model of mixture of Gaussian regressions in which the proportions are modeled using logistic weights. Using maximum likelihood estimators, a model selection procedure has been applied, supported by a theoretical guarantee. Numerical experiments have been conducted for regression mixtures with parametric logistic weights, using EM and Newton algorithms. This work is published in Electronic Journal of Statistics.

Another subject considered by Erwan Le Pennec and Lucie Montuelle was the obtention of oracle inequalities in deviation for model selection aggregation in the fixed design regression framework. Exponential weights are widely used but sub-optimal. They aggregate linear estimators and penalize Stein's unbiased risk estimate used in exponential weights to derive such inequalities. Furthermore if the infinity norm of the regression function is known and taken into account in the penalty, then a sharp oracle inequality is available. Pac-Bayesian tools and concentration inequalities play a key role in this work. These results may be found in a prepublication on arxiv or in Lucie Montuelle's PhD thesis.

In collaboration with Sylvain Arlot, Matthieu Lerasle an Patricia Reynaud-Bourret (CNRS) Nelo Molter Magalhaes considers estimator selection problem with the L^2 loss. They provide a theoretical minimal and optimal penalty. They define practical cross-validation procedures and provide non-asymptotic and first order optimal results for these procedures.

Emilie Devijver and Pascal Massart focused on the Lasso for high dimension finite mixture regression models. An ℓ_1 oracle inequality have been get for this estimator for this model, for a specific regularization parameter. Moreover, for maximum likelihood estimators, restricted to relevant variables and to low rank, theoretical results have been proved to support methodology.

Pascal Massart and Clément Levrard continue their work on the properties of the k-means algorithm in collaboration with Gérard Biau (Université Paris 6). Most of the work achieved this year was devoted to the obtention of fast convergence rates for the k-means quantizer of a source distribution in the high-dimensional case. It has been proved that the margin condition for vector quantization introduced last year can be extended to the infinite dimensional Hilbert case, and that this condition is sufficient for the source distribution to satisfy some natural properties, such as the finiteness of the set of optimal quantizers. When this condition is satisfied, a dimension-free fast convergence rate can be derived. In addition, this margin condition provides theoretical guarantees for methods combining k-means and variable selection through a Lasso-type procedure. Its implementation is still in process, however early experiments shows that this procedure can retrieve active variables in the Gaussian mixture case.

Among selection methods for nonparametric estimators, a recent one is the procedure of Goldenshluger-Lespki. This method proposes a data-driven choice of m to select an estimator among a collection $(\widehat{s}_m)_{m\in M}$. The selected $\widehat{m}$ is chosen as a minimiser of $B(m)+V(m)$ where $B(m)=\sup\{[\|\widehat{s}_m-\widehat{s}_{m'}\|-V(m')]_+,\ m'\in M\}$ and $V(m)$ is a penalty term to be suitably chosen. Previous results have established oracle inequalities to ensure that if $V(m)$ is large enough the final estimator $\widehat{s}_{\widehat{m}}$ is almost as efficient as the best one in the collection. The aim of the work of Claire Lacour and Pascal Massart was to give a practical way to calibrate $V(m)$. To do this they have evidenced an explosion phenomenon: if V is chosen smaller than some critical V_0, the risk $\|s-\widehat{s}_{\widehat{m}}\|$ is proven to dramatically increase, though for $V>V_0$ this risk is quasi-optimal. Simulations have corroborated this behavior.

The well-documented and consistent variable selection procedure in model-based cluster analysis and classification, that Cathy Maugis (INSA Toulouse) has designed during her PhD. thesis in SELECT, makes use of stepwise algorithms which are painfully slow in high dimensions. In order to circumvent this drawback, Gilles Celeux in collaboration with Mohammed Sedki (Université Paris XI) and Cathy Maugis), proposed to sort the variables using a lasso-like penalization adapted to the Gaussian mixture model context. Using this rank to select the variables they avoid the combinatory problem of stepwise procedures. After tests on challenging simulated and real data sets, their algorithm finalised and show good performances.

In collaboration with Jean-Michel Marin (Université de Montpellier) and Olivier Gascuel (LIRMM), Gilles Celeux has continued a research aiming to select a short list of models rather a single model. This short list of models is declared to be compatible with the data using a p-value derived from the Kullback-Leibler distance between the model and the empirical distribution. And, the Kullback-Leibler distances at hand are estimated through non parametric and parametric bootstrap procedures.

6.2. Statistical learning methodology and theory

Participants: Vincent Brault, Gilles Celeux, Christine Keribin, Erwan Le Pennec, Lucie Montuelle, Michel Prenat, Solenne Thivin.

Vincent Brault, Ph D. student of Gilles Celeux and Christine Keribin defended his thesis on the Latent Block Model (LBM) for categorical data. Their work investigated a Gibbs algorithm to avoid solutions with empty clusters on synthetic as well as real data (Congressional Voting Records and genomic data. They detailed the link between the information criteria ICL and BIC, compared them on synthetic and real data, and conjectured that these criteria are both consistent for LBM, which is not a standard behavior. Hence, ICL has to be preferred for LBM. This work is now published in *Statistics and Computing*.

Vincent Brault has achieved a detailed bibliographical review on coclustering with Aurore Lomet (UTC) which is currently under revision. He has also worked in collaboration with Mahindra Mariadassou (INRA) to overview the state of the art on theoretical results for latent or stochastic block model.

Vincent Brault, Christine Keribin and Mahindra Mariadassou have started a collaboration to tackle the consistency and asymptotic normality for the maximum likelihood and variational estimators in a stochastic or latent block model.

Gilles Celeux has started a collaboration with Jean-Patrick Baudry on strategies to avoid the traps of the EM algorithm in mixture analysis. They anayse the effect of the spurious local maximizers and the regulariszed algorithms to avoid these spurious solutions. They explore the link of the degree of regularization and the slope heuristics. Moreover, they propose and study strategies to initiate the EM algorithm embedding the solution with K components and the starting position with $K+1$ component to avoid suboptimal solutions.

Erwan Le Pennec is supervising Solenne Thivin in her CIFRE with Michel Prenat and Thales Optronique. The aim is target detection on complex background such as clouds or sea. Their approach is a local approach based on test decision theory. They have obtained theoretical and numerical results on a segmentation based approach in which a simple Markov field testing procedure is used in each cell of a data driven partition. They also have obtained experimental results on images (or patches) unsupervised classification, with the aim of better calibrate the detection procedure. The classification is based on features which are defined in cloud texture modeling activity.

Erwan Le Pennec and Michel Prenat have also collaborated on a cloud texture modeling using a non-parametric approach. Such a modeling coud be used to better calibrate the detection procedure: it can lead to more examples than the one acquired and it could be the basis of an ensemble method.

6.3. Reliability

Participants: Yves Auffray, Gilles Celeux, Rémy Fouchereau, Patrick Pamphile, Jana Kalawoun.

In 2014, in the framework of a CIFRE convention with Snecma-SAFRAN Rémy Fouchereau has defended a thesis on the modeling of fatigue lifetime supervised by Gilles Celeux and Patrick Pamphile. In aircraft, space and nuclear industry, fatigue test is the main basic tool for analyzing fatigue lifetime of a given material, component, or structure. A sample of the material is subjected to cyclic loading S (stress, force, strain, etc.), by a testing machine which counts N, the number of cycles to failure. Fatigue test results are plotted on a SNcurve. A probabilistic model for the construction of SN-curve is proposed. In general, fatigue test results are widely scattered for High Cycle Fatigue region and "duplex" SN-curves appears for Very High Cycle region. That is why classic models from mechanic of rupture theory on one hand, probability theory on the other hand, do not fit SN-curve on the whole range of cycles. We have proposed a probabilistic model, based on a fracture mechanic approach: few parameters are required and they are easily interpreted by mechanic or material engineers.This model has been applied to both simulated and real fatigue test data sets. The SN-curves have been well fitted on the whole range of cycles. The parameters have been estimated using the EM algorithm, combining Newton-Raphson optimisation method and Monte Carlo integral estimations. The model has been then improved taking into account production process information, thanks to a clustering approach. Thus, we have provided engineers with a probabilistic tool for reliability design of mechanical parts, but also with a diagnostic tool for material elaboration.

Since two years SELECT collaborates with CEA for the estimation of the battery State of Charge (SoC). For vehicles powered by an electric motor, SoC estimation is essential to guarantee vehicle autonomy, as well as safe utilization. The aim is to create a reliable SoC model to closely fit the battery dynamic, in embedded applications (e.g. Electric Vehicle). Jana Kalawoun started a thesis supervised by Gilles Celeux, Patrick Pamphile and Maxime Montaru (CEA) on this topic. The SoC is modeled by a Switching Markov State-Space Model. The parameters are estimated by combining the EM algorithm and Particle Filter methods. The model is validated using real-life electric vehicle data. It has been proved to be highly superior to a simple state space model. The optimal number of battery modes is then identified, using different model selection criteria as BIC or the slope heuristics.

Yves Auffray and Gilles Celeux proposed a solution to a reliability problem on Dassault's F7X business jet brakes. As the origin brake version showed poor reliability performance, an increased frequency inspection of the brakes had been decided and, after a while, a new brake version adopted. The new version has not shown any failure since its adoption. Then the question was : is it possible to relax the brakes inspection frequency ?
On the basis of first brake version failure data, the parameters of a Weibull law was estimated : $\eta = 3169, \beta = 1.38$. Under the hypothesis that the new brake version would follow the same Weibull law, the probability that none of them broke was $1.67\ 10^{-6}$. This led to reject that hypothesis.
A Weibull model for the new brakes was then estimated. The shape parameter beeing leaved conservatively unchanged, the scale parameter was estimated so that the no failure event probability amounts to 0.05. This led to $\eta = 9326$.
From the resulting Weibull model, dates $D_0, D_1, \cdots, D_k, \cdots$ of inspection for the new brakes was established so that : $\mathbb{P}(T \leq D_0 + D_1 + \cdots + D_k | T > D_0 + \cdots + D_{k-1}) = 0.01$.
Dassault has adopted this far less constraining inspection calendar.

6.4. Statistical analysis of genomic data

Participants: Vincent Brault, Gilles Celeux, Mélina Gallopin, Christine Keribin, Yann Vasseur.

In collaboration with Florence Jaffrezic and Andrea Rau (INRA, animal genetic department), Mélina Gallopin is a third year PhD student under the supervision of Gilles Celeux. This thesis is concerned with the modelization and model selection in the analysis of RNA-seq data. This year, they proposed a model selection criterion for model-based clustering of annotated gene expression data. This criterion is a ICL-like criterion taking into account the annotations. They are also working on a objective comparison of discrete and continuous modelling after a transformations for RNA-seq data based on a comparison of the likelihoods (eventually penalized) of the models in competition.

The subject of Yann Vasseur PhD Thesis, supervised by Gilles Celeux and Marie-Laure Martin-Magniette (INRA URGV), is the inference of a regulatory network on Transcriptions Factors (TFs), which are specific genes, of *Arabidopsis thaliana*. In that purpose, a transciptome dataset with a sensibly equal size of TFS and statistical units is available. The first aim consists of reducing the dimension of the network to avoid high dimension difficulties. Representing this network with a Gaussian Graphical Model, the following procedure has been defined:

1. *Selection step*: choosing the set of TFs regulators (supports) of each TF.
2. *Classification step*: deducing co-factors groups (TFs with similary expression levels) from these supports.

Thus, the reduced network would be built on the co-factors groups. Currently, several selection methods based on Gauss-LASSO and resampling procedures have been applied on the dataset. The study of the stability and the parameters calibration of these methods are in progress. The TFs are clustered with the Latent Block Model in a number of co-factors groups selected with the BIC or the exact ICL criterion.

In collaboration with Marie-Laure Martin-Magniette, Cathy Maugis and Andrea Rau, Gilles Celeux studied gene expression gotten from high-throughput sequencing technology. They focus on the question of clustering digital gene expression profiles as a means to discover groups of co-expressed genes. They propose a Poisson mixture model using a rigorous framework for parameter estimation as well as the choice of the appropriate number of clusters. They illustrate co-expression analyses using this approach on two real RNA-seq datasets. A set of simulation studies also compares the performance of the proposed model with that of several related approaches developed to cluster RNA-seq or serial analysis of gene expression data. The proposed method is implemented in the open-source R package `HTSCluster`, available on CRAN.

6.5. Model based-clustering for pharmacovigilance data

Participants: Gilles Celeux, Christine Keribin, Valérie Robert.

In collaboration with Pascale Tubert-Bitter, Ismael Ahmed and Mohamed Sedki, Gilles Celeux and Christine Keribin has started a research concerning the detection of associations between drugs and adverse events in the framework of the PhD of Valerie Robert. At first, this team has developed a model-based clustering inspired of the latent black model which consists in co-clustering rows and columns of two binary tables imposing the same row ranking. Then it enables to highlight subgroups of individuals sharing the same drug profile and subgroups of adverse effects and drugs with strong interaction. Besides, some sufficient conditions are provided to obtain the identifiability of the model and some studies are experimented on simulated data.

6.6. Curves classification, denoising and forecasting

Participants: Émilie Devijver, Pascal Massart, Jean-Michel Poggi, Vincent Thouvenot.

In collaboration with Farouk Mhamdi and Meriem Jaidane (ENIT, Tunis, Tunisia), Jean-Michel Poggi proposed a method for trend extraction from seasonal time series through the Empirical Mode Decomposition (EMD). Experimental comparison of trend extraction based on EMD, X11, X12 and Hodrick Prescott filter are conducted. First results show the eligibility of the blind EMD trend extraction method. Tunisian real peak load is also used to illustrate the extraction of the intrinsic trend.

Jean-Michel Poggi, co-supervising with Anestis Antoniadis (Université Joseph Fourier Grenoble) the PhD thesis of Vincent Thouvenot, funded by a CIFRE with EDF. The industrial motivation of this work is the recent development of new technologies for measuring power consumption by EDF to acquire consumption data for different mesh network. The thesis will focus on the development of new statistical methods for predicting power consumption by exploiting the different levels of aggregation of network data collection. From the mathematical point of view, the work is to develop generalized additive models for this type of kind of aggregated data for the modeling of functional data, associating closely nonparametric estimation and variable selection using various penalization methods.

Jean-Michel Poggi and Pascal Massart are the co-advisors of the PhD thesis of Émilie Devijver, strongly motivated by the same kind of industrial forecasting problems in electricity, which is dedicated to curves clustering for the prediction. A natural framework to explore this question is mixture of regression models for functional data. They extend to functional data the recent work by Bühlmann and coauthors dealing with the simultaneous estimation of mixture regression models in the scalar case using Lasso type methods. It is based on the technical tools of the work of Caroline Meynet (which completes her thesis Orsay under the direction of P. Massart), which deals with the clustering of functional data using Lasso methods choosing simultaneously number of clusters and selecting significant wavelet coefficients. Nevertheless, they also propose a procedure dealing with low rank estimator. Simulations and benchmark data have been conducted for high-dimensional finite mixture regression models.

Jean-Michel Poggi, co-supervising with Meriem Jaëdane, Raja Ghozi (ENIT Tunisie) and from the industrial side, Sylvie Sevestre-Ghalila (CEA LinkLab), the PhD thesis of Neska El Haouij, funded by a kind of CIFRE with CEA LinkLab. The industrial motivation of this work is the recent development of new technologies for sensory measurements, environmental and physiological to explain and improve the driving tasks . The thesis aims to explain sensory aspects involved in automated decision to the car interior, by objectivization. The thesis will focus on the use and development of experimental designs and statistical methods to quantify and explain driving ability in to the modeling using functional explanatory factors. Statistical contributions of this work will involve nonparametric estimation and variable selection and/or models.

6.7. Statistical analysis of medical images

Participants: Christine Keribin, Yves Rozenholc.

Yves Rozenholc and C. Keribin work the genomic tumoral alterations and supervised a Master student Yi LIU. The study of genomic DNA alterations (recurrent regions of alteration, patterns of instability) contributes to tumor classification, and becomes of great importance for the personalization of cancer treatments. The use of Single-Nucleotide Polymorphism (SNP) arrays or of New Generation Sequences (NGS) techniques allows the simultaneous estimation of segmented copy number (CN) and B-allele frequency (BAF) profiles along the whole genome. In this context, Popova (2009) proposed the GAP method, based on pattern recognition with (BAF, CN) maps to detect genotype status of each segment in complex tumoral genome profiles. It takes into account the fact that the observations on these maps are necessarily placed on centers that depend –up to a proper scaling of the CN– only on the unknown proportion of non tumoral tissue in the sample. Being deterministic and manually tuned, this method appears sensitive to noise. To overcome this drawback, they set a mixture model, allowing the automatic estimation of the proportion of non tumoral tissue and the test of genotype for each segment along the whole genome. They develop the estimation with an adapted EM algorithm that has been tested on simulated data. This work has already been presented (ERCIM 14, SEQBIO14) and provides many potential developments.

SEMAGRAMME Project-Team

6. New Results

6.1. Generation

G-TAG [52], [66] is a Tree Adjoining Grammar (TAG) based formalism which was specifically designed for the task of text generation. Contrary to TAG, the derivation structure becomes primary, as pivot between the conceptual representation and the surface form. This is a shared feature with the encoding of TAG into ACG. Laurence Danlos (Alpage Inria project), Aleksandre Maskharashvili, and Sylvain Pogodalla have shown how to recast the G-TAG formalism into ACG, relying on the reversibility properties of the later [17], [16], [18].

6.2. Discourse Grammar

Laurence Danlos (Alpage Inria project), Aleksandre Maskharashvili, and Sylvain Pogodalla have presented a method to interface a sentential grammar and a discourse grammar. It offers both a smooth integration of the two grammars without using an intermediate processing step, and the possibility to build discourse structures that are direct acyclic graphs (DAG) and not only trees. The analysis is based on a Tree-Adjoining Grammar (TAG) approach to discourse: Discourse Synchronous TAG (D-STAG) [50], [51], and uses an encoding of TAG into ACG. This allows for expressing a higher-order semantic interpretation that enables building DAG discourse structures, and for smoothly integrating the sentential and the discourse grammar thanks to the modular capability of ACG. All the examples may be run and tested with the the ACGtk (submitted).

6.3. Large Scale Grammatical Resources

Guy Perrier wrote a complete documentation [36] on FRIGRAM [0] a French grammar with a large coverage, written in the formalism of Interaction Grammars [59]. The different chapters of the 257 pages of documentation correspond to the different parts of speech in French. At the end, two chapters are dedicated to two specific phenomena: extraction (relative, interrogative and cleft clauses) and coordination, which is presented in common with punctuation because of their proximity.

6.4. Deep Syntax Annotation of the Sequoia French Treebank

Marie Candito, Guy Perrier, Bruno Guillaume, Corentin Ribeyre, Karën Fort, Djamé Seddah and Eric de la Clergerie annotated the Sequoia French Treebank with deep syntax dependencies [14].

The Sequoia French Treebank [47] is a 3.100 sentences treebank covering several domains (news, medical, europarl and fr-wikipedia). It is freely available and has already been annotated with surface dependency representations.

The participants in the project have defined a deep syntactic representation scheme for French, built from the surface annotation scheme of the Sequoia corpus and abstracting away from it [28]. This scheme expresses the grammatical relations between content words. When these grammatical relations take part into verbal diatheses, the diatheses are considered as resulting from redistributions from the canonical diathesis, which is retained in the annotation scheme.

The goal is to obtain a freely available corpus, which will be useful for corpus linguistics studies and for training deep analyzers to prepare semantic analysis.

The different steps of the annotation process were conducted in a collaborative way. As the members of the project are located in two different French towns (Paris and Nancy), they decided to produce a complete annotation of the TreeBank in both towns and to collaboratively adjudicate the two results.

[0] http://wikilligramme.loria.fr/doku.php/frig:frig

Each team separately produced an initial annotated version of the mini reference. The final version, resulting from several iterations and adjudications, is available [0].

6.5. Exploitation of the LVF (Lexicon of French Verbs)

Bruno Guillaume, Karën Fort, Guy Perrier and Paul Bédaride have worked on the LVF [53] ("Lexique des Verbes du Français", Lexicon of French Verbs). This large lexicon was build by two French linguists, Jean Dubois and Françoise Dubois-Charlier and contains detailed linguistic information about 12.308 lemmas of French verbs. The work presented in [21] describes experiments aiming at mapping the LVF to DICOVALENCE [87]. The two resources (LVF and DICOVALENCE) were built by linguists, based on very different theories, which makes a direct mapping nearly impossible. In the current work, we focus on the linguistic examples given in LVF. These examples are not sentences that can be parsed directly; the first part of the work was to express examples as real natural language sentence. It is then possible to use FRILEX, a Natural Language Processing lexicon based on DICOVALENCE to parse corrected examples given with LVF entries. This results in an automatic partial mapping of LVF entries against DICOVALENCE entries.

6.6. Game With A Purpose

Crowdsourcing is nowadays a way of constructing linguistic resources which is more and more used. In the crowdsourcing area, one of the way to motivate a large amount of people to contribute to a project is to present it as a game. Games used in this particular way are called GWAPs (Game With A Purpose).

In Natural Language Processing, examples of GWAP are "Phrase detective" where games are asked to resolve anaphora in English texts and "JeuDeMots" where gamers have to given lexical terms related to a term given by the system (the goal is to build a semantic networks of French lexical items).

Karën Fort and Bruno Guillaume worked on the definition of a GWAP to help construction of syntactically annotated corpora. With a student (Hadrien Chastant), they presented in April, the design of ZombiLingo [20], a GWAP that allows for the dependency syntax annotation of French corpora. The main aspects of this work are to explain: how to deal with the complexity of the task, how to motivate gamers to contribute and how to ensure that a large numbers of gamers will help to produce an high quality linguistic resource.

With another student (Valentin Stern), a first prototype was built. This first version implements only a part of the mechanisms described in the previous work and it is used as a proof-of-concept of a future game. This prototype was presentend at the TALN conference in July [27].

6.7. Supertagging

Guillaume Bonfante, Bruno Guillaume, Mathieu Porey and Guy Perrier wrote a book chapter [30] "Suppertagging with constraints". This chapter makes a survey of the results obtained in previous publications about supertagging based on polarities [46] and based on the companionship principle [45]. The last section of the chapter presents a new application of the companionship principle to the TAG formalism and presents some experimental results.

6.8. Modelling Semantic Phenomena

Despite the valuable insights yielded by the classical theories of discourse semantics, there is a wide range of exceptional phenomena that they fail to address, e.g., anaphora under double negation and modality. Concentrating on these two exceptions, Sai Qian, Philippe de Groote and Maxime Amblard provide a corresponding adaptation of TTDL for each case. Briefly speaking, for the problem of double negation, they propose to encapsulate both the affirmative representation and the negative representation of an expression in its semantics. Negation is treated as an operation which switches the positions of the two representations. Thus a second negation will switch the positions again as if no negation had ever occurred. In this way, a double negation can be eliminated and the desired referent accessibility is modelled. As for anaphora under modality,

[0] https://deep-sequoia.inria.fr

they propose to enrich the TTDL left context with the notion of modal base, which is proposed by Kratzer. The possible world model is integrated in the semantic representation as well. Moreover, they show how the different adaptations could work in an unified framework, [75].

6.9. Quantification in event semantics

Yoad Winter (Utrecht University) has given a type-logical account of quantification in event semantics.

It has been observed in the literature that Davidson's event semantics does not combine smoothly with Montague's compositional semantics. The difficulty comes from a possibly bad interaction between event existential closure, on the one hand, and quantification, negation, or conjunction, on the other hand. In a recent publication, Winter and Zwarts provide a solution to this problem. Winter and de Groote elaborate on this solution. In particular, they provide a treatment of quantified adverbial modifiers, which was absent from Winter and Zwarts, [19].

6.10. Pragmasemantic with Effects and Handlers

Jiří Maršík and Maxime Amblard have explored the feasibility of theories of side effects of programming languages in the study of natural language semantics and pragmatics [23]. In the approach that we are developing, the denotations we assign to fragments natural language are effectful computations. To demonstrate on an example, if we was to treat dynamics, then instead of changing the type of sentence denotations from o to c -> o * c, where c is the type of discourse contexts and o is the type of propositions, we would treat sentence denotations as effectful computations of type o that study and modify the context using effectful operations. This explicit distinction between 'result' and 'effects' brings to mind Stalnaker's distinction between 'content' and 'context'.

The motivation for this approach is to make it easier to compose multiple pragmasemantic phenomena by being allowed to put their effects aside. So far, a small prototype handling dynamics, presuppositions and some of their interactions is under development.

6.11. Mining Texts at discourse level

Linguistic discourse refers to the meaning of large chunks of text, from phrases to whole documents. It could be very useful for guiding attempts at text mining, which focus on document selection, document summarization, or other knowledge extraction goals. Hence the aim of this work is to apply Knowledge Discovery in Databases (KDD) methods to texts annotated with discourse information. Maxime Amblard with Yannick Toussaint (Orpailleur team) and Sara van de Moosdijk (master 2 intern) approach the problem by extracting discourse relations using unsupervised methods, which are then used to construct a knowledge model with Formal Concept Analysis (FCA). Pattern Structures (PS), an advancement in FCA, allow for the modelling of complex data. Our method is applied to a corpus of medical articles compiled from PubMed. This medical data is enhanced with concepts from the UMLS MetaThesaurus combined with the UMLS Semantic Network to serve as an ontology for Pattern Structure classification. The results show that despite having a large amount of noise, the method is promising and could be applied to other domains than the medical domain. We explore the pitfalls and suggest ways in which the process could be improved (Submission under review).

6.12. Exploring real datas

Maxime Amblard explored the use of formal framework for modelling transcription of real interviews, in particular one involves in the SLAM project with schizophrenics. Schizophrenia is well-known among mental illnesses for the severity of the thought disorders it involves, and for their widespread and spectacular manifestations: from deviant social behavior to delusion, not to mention affective and sensitive distortions. The goal of our interdisciplinary work is to (i) analyze linguistic troubles in conversational contexts in which one of the speakers is schizophrenic, (ii) construe how the concept of rationality and logicality may apply to them, and (iii) propose a formal representation about this specific manifestation. Maxime Amblard, Sylvain Pogodalla and Karen Fort propose surveys on past results [35], [29].

Maxime Amblard and Karen Fort have studied experiments they led concerning disfluencies in the discourse of schizophrenic patients (in remediation). These experiments are part of a larger study dealing with other levels of linguistic analysis, that could eventually help identifying clues leading to the diagnostic of the disease. This study largely relies on natural language processing tools, which allow for the rapid processing of massive textual data (here, more than 375,000 words). The first phase of the study, which they present confirmed the correlation between schizophrenia and the number of disfluences appearing in the discourse [25]. Moreover they have discussed ethical issues on the corpus with others [26].

6.13. Paraconsistency and Inconsistency-Friendly Logics

Paraconsistent logic is a family of formal systems in which the law of contradiction fails. In such systems, from an inconsistent set, *not* everything follows.

Can Baskent has studied such logical systems and their connections to formal linguistics within the framework of game theory. First, he observed how a game theoretical semantics can be given for some paraconsistent logics [43] . The advantage of game semantics is that it simply reflects the parsing tree of logic, and furthermore presents a semantical structure that uses elements from game theory. Such a study also requires an in-depth study of various paraconsistent logics, and their semantical structures [13]. Such a study requires some understanding of point-set topology, and its relation to logic.

Moreover, paraconsistent logics relate to dynamic logics as well. The logical model defines characterises how dynamic epistemic modalities, which are familiar from multi-agent systems, work [13]. This helps us understand how multi-agent interactions in an inconsistent model work in a sound way.

Another interesting way of seeing how inconsistency-friendly logics work is to consider them within the framework of game theory [37]. Game theory, similar to multi-agent systems, studies the intelligent and rational interaction of decision makers/agents. Yet, it suffers from various paradoxes. Such paradoxes are important from a computational semantical point of view. If paraconsistency is the most suitable tool to analyse paradoxes, then game theoretical paradoxes are not exceptions [37].

The technical work always needs to be supplemented by some conceptual work. Granted, paraconsistent logics find their ways in various philosophical and semantical issues, yet their computational analysis usually falls short. In [44], we discussed the connection between paraconsistent logics and Hintikka's interrogative models. These models have been developed by Hintikka, a pioneer of epistemic logic, and have been properly analysed from paraconsistent perspectives. If inquiry and questioning needs to be accounted for computationally, a paraconsistent approach will be an appropriate tool as well. Similarly, [39] discusses paraconsistency and its connection to social software. *Social Software* is a field conceived by Rohit Parikh, and it studies the computational and logical analysis of social protocols and policies. It lies in the intersection of social choice theory and game theory, and is a subset of logic.

Such results have been presented in various talks including, *World Congress of Paraconsistency* in Kolkata and *Logic and the Foundations of Game and Decision Theory* in Bergen, and warmly received.

SEQUEL Project-Team

6. New Results

6.1. Highlights of the Year

- New startup by Rémi Coulom on AI in games (go, chess, mahjong, $\cdots$).
- Successful Collaboration with Deezer and the victory at the ACM RecSys Recommendation Systems Challenge
- We were selected and working on preparation of ICML 2015 in Lille. ICML is the most important conference in the field of machine learning. This is the first time after more than 30 years of existence, that this conference will be held in France.

6.2. Decision-making Under Uncertainty

6.2.1. Reinforcement Learning

Selecting Near-Optimal Approximate State Representations in Reinforcement Learning [23]

We consider a reinforcement learning setting where the learner does not have explicit access to the states of the underlying Markov decision process (MDP). Instead, she has access to several models that map histories of past interactions to states. Here we improve over known regret bounds in this setting, and more importantly generalize to the case where the models given to the learner do not contain a true model resulting in an MDP representation but only approximations of it. We also give improved error bounds for state aggregation.

Online Stochastic Optimization under Correlated Bandit Feedback [15]

In this paper we consider the problem of online stochastic optimization of a locally smooth function under bandit feedback. We introduce the high-confidence tree (HCT) algorithm, a novel anytime X -armed bandit algorithm, and derive regret bounds matching the performance of state-of-the-art algorithms in terms of the dependency on number of steps and the near-optimality di-mension. The main advantage of HCT is that it handles the challenging case of correlated bandit feedback (reward), whereas existing meth-ods require rewards to be conditionally indepen-dent. HCT also improves on the state-of-the-art in terms of the memory requirement, as well as requiring a weaker smoothness assumption on the mean-reward function in comparison with the existing anytime algorithms. Finally, we discuss how HCT can be applied to the problem of policy search in reinforcement learning and we report preliminary empirical results.

Sparse Multi-task Reinforcement Learning [9]

In multi-task reinforcement learning (MTRL), the objective is to simultaneously learn multiple tasks and exploit their similarity to improve the performance w.r.t. single-task learning. In this paper we investigate the case when all the tasks can be accurately represented in a linear approximation space using the same small subset of the original (large) set of features. This is equivalent to assuming that the weight vectors of the task value functions are *jointly sparse*, i.e., the set of their non-zero components is small and it is shared across tasks. Building on existing results in multi-task regression, we develop two multi-task extensions of the fitted Q-iteration algorithm. While the first algorithm assumes that the tasks are jointly sparse in the given representation, the second one learns a transformation of the features in the attempt of finding a more sparse representation. For both algorithms we provide a sample complexity analysis and numerical simulations.

6.2.2. Multi-arm Bandit Theory

Spectral Bandits for Smooth Graph Functions with Applications in Recommender Systems [20]

Smooth functions on graphs have wide applications in manifold and semi-supervised learning. In this paper, we study a bandit problem where the payoffs of arms are smooth on a graph. This framework is suitable for solving online learning problems that involve graphs, such as content-based recommendation. In this problem, each recommended item is a node and its expected rating is similar to its neighbors. The goal is to recommend items that have high expected ratings. We aim for the algorithms where the cumulative regret would not scale poorly with the number of nodes. In particular, we introduce the notion of an effective dimension, which is small in real-world graphs, and propose two algorithms for solving our problem that scale linearly in this dimension. Our experiments on real-world content recommendation problem show that a good estimator of user preferences for thousands of items can be learned from just tens nodes evaluations.

Online combinatorial optimization with stochastic decision sets and adversarial losses [21]

Most work on sequential learning assumes a fixed set of actions that are available all the time. However, in practice, actions can consist of picking subsets of readings from sensors that may break from time to time, road segments that can be blocked or goods that are out of stock. In this paper we study learning algorithms that are able to deal with stochastic availability of such unreliable composite actions. We propose and analyze algorithms based on the Follow-The-Perturbed-Leader prediction method for several learning settings differing in the feedback provided to the learner. Our algorithms rely on a novel loss estimation technique that we call Counting Asleep Times. We deliver regret bounds for our algorithms for the previously studied full information and (semi-)bandit settings, as well as a natural middle point between the two that we call the restricted information setting. A special consequence of our results is a significant improvement of the best known performance guarantees achieved by an efficient algorithm for the sleeping bandit problem with stochastic availability. Finally, we evaluate our algorithms empirically and show their improvement over the known approaches.

Extreme bandits [10]

In many areas of medicine, security, and life sciences, we want to allocate limited resources to different sources in order to detect extreme values. In this paper, we study an efficient way to allocate these resources sequentially under limited feedback. While sequential design of experiments is well studied in bandit theory, the most commonly optimized property is the regret with respect to the maximum mean reward. However, in other problems such as network intrusion detection, we are interested in detecting the most extreme value output by the sources. Therefore, in our work we study extreme regret which measures the efficiency of an algorithm compared to the oracle policy selecting the source with the heaviest tail. We propose the ExtremeHunter algorithm, provide its analysis, and evaluate it empirically on synthetic and real-world experiments.

Efficient learning by implicit exploration in bandit problems with side observations [18]

We consider online learning problems under a a partial observability model capturing situations where the information conveyed to the learner is between full information and bandit feedback. In the simplest variant, we assume that in addition to its own loss, the learner also gets to observe losses of some other actions. The revealed losses depend on the learner's action and a directed observation system chosen by the environment. For this setting, we propose the first algorithm that enjoys near-optimal regret guarantees without having to know the observation system before selecting its actions. Along similar lines, we also define a new partial information setting that models online combinatorial optimization problems where the feedback received by the learner is between semi-bandit and full feedback. As the predictions of our first algorithm cannot be always computed efficiently in this setting, we propose another algorithm with similar properties and with the benefit of always being computationally efficient, at the price of a slightly more complicated tuning mechanism. Both algorithms rely on a novel exploration strategy called implicit exploration, which is shown to be more efficient both computationally and information-theoretically than previously studied exploration strategies for the problem.

Best-Arm Identification in Linear Bandits [29]

We study the best-arm identification problem in linear bandit, where the rewards of the arms depend linearly on an unknown parameter θ^* and the objective is to return the arm with the largest reward. We characterize the complexity of the problem and introduce sample allocation strategies that pull arms to identify the best arm with a fixed confidence, while minimizing the sample budget. In particular, we show the importance of exploiting the global linear structure to improve the estimate of the reward of near-optimal arms. We analyze the proposed strategies and compare their empirical performance. Finally, we point out the connection to the G-optimality criterion used in optimal experimental design.

Exploiting easy data in online optimization [28]

We consider the problem of online optimization, where a learner chooses a decision from a given decision set and suffers some loss associated with the decision and the state of the environment. The learner's objective is to minimize its cumulative regret against the best fixed decision in hindsight. Over the past few decades numerous variants have been considered, with many algorithms designed to achieve sub-linear regret in the worst case. However, this level of robustness comes at a cost. Proposed algorithms are often over-conservative, failing to adapt to the actual complexity of the loss sequence which is often far from the worst case. In this paper we introduce a general algorithm that, provided with a "safe" learning algorithm and an opportunistic "benchmark", can effectively combine good worst-case guarantees with much improved performance on "easy" data. We derive general theoretical bounds on the regret of the proposed algorithm and discuss its implementation in a wide range of applications, notably in the problem of learning with shifting experts (a recent COLT open problem). Finally, we provide numerical simulations in the setting of prediction with expert advice with comparisons to the state of the art.

Spectral Bandits for Smooth Graph Functions [32]

Smooth functions on graphs have wide applications in manifold and semi-supervised learning. In this paper, we study a bandit problem where the payoffs of arms are smooth on a graph. This framework is suitable for solving online learning problems that involve graphs, such as content-based recommendation. In this problem, each item we can recommend is a node and its expected rating is similar to its neighbors. The goal is to recommend items that have high expected ratings. We aim for the algorithms where the cumulative regret with respect to the optimal policy would not scale poorly with the number of nodes. In particular, we introduce the notion of an effective dimension, which is small in real-world graphs, and propose two algorithms for solving our problem that scale linearly and sublinearly in this dimension. Our experiments on real-world content recommendation problem show that a good estimator of user preferences for thousands of items can be learned from just tens of nodes evaluations.

Regret bounds for restless Markov bandits [5]

We consider the restless Markov bandit problem, in which the state of each arm evolves according to a Markov process independently of the learner's actions. We suggest an algorithm, that first represents the setting as an MDP which exhibits some special structural properties. In order to grasp this information we introduce the notion of ϵ-structured MDPs, which are a generalization of concepts like (approximate) state aggregation and MDP homomorphisms. We propose a general algorithm for learning ϵ-structured MDPs and show regret bounds that demonstrate that additional structural information enhances learning. Applied to the restless bandit setting, this algorithm achieves after any T steps regret of order $\widetilde{O}(T^{1/2})$ with respect to the best policy that knows the distributions of all arms. We make no assumptions on the Markov chains underlying each arm except that they are irreducible. In addition, we show that index-based policies are necessarily suboptimal for the considered problem.

Spectral Thompson Sampling [19]

Thompson Sampling (TS) has surged a lot of interest due to its good empirical performance, in particular in the computational advertising. Though successful, the tools for its performance analysis appeared only recently. In this paper, we describe and analyze SpectralTS algorithm for a bandit problem, where the payoffs of the choices are smooth given an underlying graph. In this setting, each choice is a node of a graph and the expected payoffs of the neighboring nodes are assumed to be similar. Although the setting has application both in recommender systems and advertising, the traditional algorithms would scale poorly with the number of choices. For that purpose we consider an effective dimension d, which is small in real-world graphs. We deliver the analysis showing that the regret of SpectralTS scales as $d(T\ln N)^{1/2}$ with high probability, where T is the time horizon and N is the number of choices. Since a d sqrt(T ln N) regret is comparable to the known results, SpectralTS offers a computationally more efficient alternative. We also show that our algorithm is competitive on both synthetic and real-world data.

6.2.3. Recommendation systems

User Engagement as Evaluation: a Ranking or a Regression Problem? [39]

In this paper, we describe the winning approach used on the RecSys Challenge 2014 which focuses on employing user en-gagement as evaluation of recommendations. On one hand, we regard the challenge as a ranking problem and apply the LambdaMART algorithm, which is a listwise model special-ized in a Learning To Rank approach. On the other hand, after noticing some specific characteristics of this challenge, we also consider it as a regression problem and use pointwise regression models such as Random Forests. We compare how these different methods can be modified or combined to improve the accuracy and robustness of our model and we draw the advantages or disadvantages of each approach.

Improving offline evaluation of contextual bandit algorithms via bootstrapping techniques [22]

In many recommendation applications such as news recommendation, the items that can be recommended come and go at a very fast pace. This is a challenge for recommender systems (RS) to face this setting. Online learning algorithms seem to be the most straight forward solution. The contextual bandit framework was introduced for that very purpose. In general the evaluation of a RS is a critical issue. Live evaluation is often avoided due to the potential loss of revenue, hence the need for offline evaluation methods. Two options are available. Model based meth- ods are biased by nature and are thus difficult to trust when used alone. Data driven methods are therefore what we consider here. Evaluat- ing online learning algorithms with past data is not simple but some methods exist in the litera- ture. Nonetheless their accuracy is not satisfac- tory mainly due to their mechanism of data re- jection that only allow the exploitation of a small fraction of the data. We precisely address this issue in this paper. After highlighting the limita- tions of the previous methods, we present a new method, based on bootstrapping techniques. This new method comes with two important improve- ments: it is much more accurate and it provides a measure of quality of its estimation. The latter is a highly desirable property in order to minimize the risks entailed by putting online a RS for the first time. We provide both theoretical and ex- perimental proofs of its superiority compared to state-of-the-art methods, as well as an analysis of the convergence of the measure of quality.

Bandits Warm-up Cold Recommender Systems [35]

We address the cold start problem in recommendation systems assuming no contextual information is available neither about users, nor items. We consider the case in which we only have access to a set of ratings of items by users. Most of the existing works consider a batch setting, and use cross-validation to tune parameters. The classical method consists in minimizing the root mean square error over a training subset of the ratings which provides a factorization of the matrix of ratings, interpreted as a latent representation of items and users. Our contribution in this paper is 5-fold. First, we explicit the issues raised by this kind of batch setting for users or items with very few ratings. Then, we propose an online setting closer to the actual use of recommender systems; this setting is inspired by the bandit framework. The proposed methodology can be used to turn any recommender system dataset (such as Netflix, MovieLens,...) into a sequential dataset. Then, we explicit a strong and insightful link between contextual bandit algorithms and matrix factorization; this leads us to a new algorithm that tackles the exploration/exploitation dilemma associated to the cold start problem in a strikingly new perspective. Finally, experimental evidence confirm that our algorithm is effective in dealing with the cold start problem on publicly available datasets. Overall, the goal of this paper is to bridge the gap between recommender systems based on matrix factorizations and those based on contextual bandits.

6.2.4. Nonparametric statistics of time series

Uniform hypothesis testing for finite-valued stationary processes [6]

Given a discrete-valued sample $X_1, \cdots, X_n$ we wish to decide whether it was generated by a distribution belonging to a family H_0, or it was generated by a distribution belonging to a family H_1. In this work we assume that all distributions are stationary ergodic, and do not make any further assumptions (e.g. no independence or mixing rate assumptions). We would like to have a test whose probability of error (both Type I and Type II) is uniformly bounded. More precisely, we require that for each ϵ there exist a sample size n such that probability of error is upper-bounded by ϵ for samples longer than n. We find some necessary and some sufficient conditions on H_0 and H_1 under which a consistent test (with this notion of consistency) exists. These conditions are topological, with respect to the topology of distributional distance.

Asymptotically consistent estimation of the number of change points in highly dependent time series [17]

The problem of change point estimation is considered in a general framework where the data are generated by arbitrary unknown stationary ergodic process distributions. This means that the data may have long-range dependencies of an arbitrary form. In this context the consistent estimation of the number of change points is provably impossible. A formulation is proposed which overcomes this obstacle: it is possible to find the correct number of change points at the expense of introducing the additional constraint that the correct number of process distributions that generate the data is provided. This additional parameter has a natural interpretation in many real-world applications. It turns out that in this formulation change point estimation can be reduced to time series clustering. Based on this reduction, an algorithm is proposed that finds the number of change points and locates the changes. This algorithm is shown to be asymptotically consistent. The theoretical results are complemented with empirical evaluations.

6.3. Statistical Learning and Bayesian Analysis

6.3.1. Prediction of Sequences of Structured and Unstructured Data

Statistical performance analysis of a fast super-resolution technique using noisy translations [38]

It is well known that the registration process is a key step for super-resolution reconstruction. In this work, we propose to use a piezoelectric system that is easily adaptable on all microscopes and telescopes for controlling accurately their motion (down to nanometers) and therefore acquiring multiple images of the same scene at different controlled positions. Then a fast super-resolution algorithm can be used for efficient super-resolution reconstruction. In this case, the optimal use of r2 images for a resolution enhancement factor r is generally not enough to obtain satisfying results due to the random inaccuracy of the positioning system. Thus we propose to take several images around each reference position. We study the error produced by the super-resolution algorithm due to spatial uncertainty as a function of the number of images per position. We obtain a lower bound on the number of images that is necessary to ensure a given error upper bound with probability higher than some desired confidence level.

Quantitative control of the error bounds of a fast super-resolution technique for microscopy and astronomy [11]

While the registration step is often problematic for super-resolution, many microscopes and telescopes are now equipped with a piezoelectric mechanical system which permits to ac-curately control their motion (down to nanometers). There-fore one can use such devices to acquire multiple images of the same scene at various controlled positions. Then a fast super-resolution algorithm [1] can be used for efficient super-resolution. However the minimal use of r 2 images for a resolution enhancement factor r is generally not sufficient to obtain good results. We propose to take several images at po-sitions randomly distributed close to each reference position. We study the number of images necessary to control the error resulting from the super-resolution algorithm by [1] due to the uncertainty on positions. The main result is a lower bound on the number of images to respect a given error upper bound with probability higher than a desired confidence level.

6.3.2. Statistical analysis of superresolution

A diffusion strategy for distributed dictionary learning [12]

We consider the problem of a set of nodes which is required to collectively learn a common dictionary from noisy measurements. This distributed dictionary learning approach may be useful in several contexts including sensor networks. Dif-fusion cooperation schemes have been proposed to estimate a consensus solution to distributed linear regression. This work proposes a diffusion-based adaptive dictionary learning strategy. Each node receives measurements which may be shared or not with its neighbors. All nodes cooperate with their neighbors by sharing their local dictionary to estimate a common representa-tion. In a diffusion approach, the resulting algorithm corresponds to a distributed alternate optimization. Beyond dictionary learn-ing, this strategy could be adapted to many matrix factorization problems in various settings. We illustrate its efficiency on some numerical experiments, including the difficult problem of blind hyperspectral images unmixing.

6.4. Miscellaneous

6.4.1. Miscellaneous

Online Matrix Completion Through Nuclear Norm Regularisation [14]

It is the main goal of this paper to propose a novel method to perform matrix completion on-line. Motivated by a wide variety of applications, ranging from the design of recommender systems to sensor network localization through seismic data reconstruction, we consider the matrix completion problem when entries of the matrix of interest are observed gradually. Precisely, we place ourselves in the situation where the predictive rule should be refined incrementally, rather than recomputed from scratch each time the sample of observed entries increases. The extension of existing matrix completion methods to the sequential prediction context is indeed a major issue in the Big Data era, and yet little addressed in the literature. The algorithm promoted in this article builds upon the Soft Impute approach introduced in Mazumder et al. (2010). The major novelty essentially arises from the use of a randomised technique for both computing and updating the Singular Value Decomposition (SVD) involved in the algorithm. Though of disarming simplicity, the method proposed turns out to be very efficient, while requiring reduced computations. Several numerical experiments based on real datasets illustrating its performance are displayed, together with preliminary results giving it a theoretical basis.

Synthèse en espace et temps du rayonnement acoustique d'une paroi sous excitation turbulente par synthèse spectrale 2D+T et formulation vibro-acoustique directe [33]

Une méthode directe pour simuler les vibrations et le rayonnement acoustique d'une paroi soumise à un écoulement subsonique est proposée. Tout d'abord, en adoptant l'hypothèse d'un écoulement homogène et stationnaire, on montre qu'une méthode de synthèse spectrale en espace et temps (2D+t) est suffisante pour obtenir explicitement une réalisation d'un champ de pression pariétale excitatrice p(x,y,t) dont les propriétés inter-spectrales sont prescrites par un modèle empirique de Chase. Cette pression turbulente p(x,y,t) est obtenue explicitement et permet de résoudre le problème vibro-acoustique de la paroi dans une formulation directe. La méthode proposée fournit ainsi une solution complète du problème dans le domaine spatio-temporel : pression excitatrice, déplacement en flexion et pression acoustique rayonnée par la paroi. Une caractéristique de la méthode proposée est un cout de calcul qui s'avère similaire aux formulations inter-spectrales majoritairement utilisées dans la littérature. En particulier, la synthèse permet de prendre en compte l'intégralité des échelles spatio-temporelles du problème : échelles turbulentes, vibratoires et acoustiques. A titre d'exemple, la pression aux oreilles d'un auditeur suite à l'excitation turbulente de la paroi est synthétisée.

Bandits attack function optimization [27]

We consider function optimization as a sequential decision making problem under the budget constraint. Such constraint limits the number of objective function evaluations allowed during the optimization. We consider an algorithm inspired by a continuous version of a multi-armed bandit problem which attacks this optimization problem by solving the tradeoff between exploration (initial quasi-uniform search of the domain) and exploitation (local optimization around the potentially global maxima). We introduce the so-called Simultaneous Optimistic Optimization (SOO), a deterministic algorithm that works by domain partitioning. The benefit of such an approach are the guarantees on the returned solution and the numerical eficiency of the algorithm. We present this machine learning rooted approach to optimization, and provide the empirical assessment of SOO on the CEC'2014 competition on single objective real-parameter numerical optimization testsuite.

Optimistic planning in Markov decision processes using a generative model [30]

We consider the problem of online planning in a Markov decision process with discounted rewards for any given initial state. We consider the PAC sample com-plexity problem of computing, with probability 1-δ, an -optimal action using the smallest possible number of calls to the generative model (which provides reward and next-state samples). We design an algorithm, called StOP (for Stochastic-Optimistic Planning), based on the "optimism in the face of uncertainty" princi-ple. StOP can be used in the general setting, requires only a generative model, and enjoys a complexity bound that only depends on the local structure of the MDP.

Near-Optimal Rates for Limited-Delay Universal Lossy Source Coding [3]

We consider the problem of limited-delay lossy coding of individual sequences. Here, the goal is to design (fixed-rate) compression schemes to minimize the normalized expected distortion redundancy relative to a reference class of coding schemes, measured as the difference between the average distortion of the algorithm and that of the best coding scheme in the reference class. In compressing a sequence of length T, the best schemes available in the literature achieve an $O(T^{-1/3})$ normalized distortion redundancy relative to finite reference classes of limited delay and limited memory, and the same redundancy is achievable, up to logarithmic factors, when the reference class is the set of scalar quantizers. It has also been shown that the distortion redundancy is at least of order $T^{-1/2}$ in the latter case, and the lower bound can easily be extended to sufficiently powerful (possibly finite) reference coding schemes. In this paper, we narrow the gap between the upper and lower bounds, and give a compression scheme whose normalized distortion redundancy is $O(ln(T)/T^{1/2})$ relative to any finite class of reference schemes, only a logarithmic factor larger than the lower bound. The method is based on the recently introduced shrinking dartboard prediction algorithm, a variant of exponentially weighted average prediction. The algorithm is also extended to the problem of joint source-channel coding over a (known) stochastic noisy channel and to the case when side information is also available to the decoder (the Wyner–Ziv setting). The same improvements are obtained for these settings as in the case of a noiseless channel. Our method is also applied to the problem of zero-delay scalar quantization, where $O(ln(T)/T^{1/2})$ normalized distortion redundancy is achieved relative to the (infinite) class of scalar quantizers of a given rate, almost achieving the known lower bound of order $1/T^{-1/2}$. The computationally efficient algorithms known for scalar quantization and the Wyner–Ziv setting carry over to our (improved) coding schemes presented in this paper.

Online Markov Decision Processes Under Bandit Feedback [4]

Software systems are composed of many interacting elements. A natural way to abstract over software systems is to model them as graphs. In this paper we consider software dependency graphs of object-oriented software and we study one topological property: the degree distribution. Based on the analysis of ten software systems written in Java, we show that there exists completely different systems that have the same degree distribution. Then, we propose a generative model of software dependency graphs which synthesizes graphs whose degree distribution is close to the empirical ones observed in real software systems. This model gives us novel insights on the potential fundamental rules of software evolution.

A Generative Model of Software Dependency Graphs to Better Understand Software Evolution [37]

Software systems are composed of many interacting elements. A natural way to abstract over software systems is to model them as graphs. In this paper we consider software dependency graphs of object-oriented software and we study one topological property: the degree distribution. Based on the analysis of ten software systems written in Java, we show that there exists completely different systems that have the same degree distribution. Then, we propose a generative model of software dependency graphs which synthesizes graphs whose degree distribution is close to the empirical ones observed in real software systems. This model gives us novel insights on the potential fundamental rules of software evolution.

Preference-Based Rank Elicitation using Statistical Models: The Case of Mallows [8]

We address the problem of rank elicitation as-suming that the underlying data generating pro-cess is characterized by a probability distribu-tion on the set of all rankings (total orders) of a given set of items. Instead of asking for complete rankings, however, our learner is only allowed to query pairwise preferences. Using information of that kind, the goal of the learner is to reliably predict properties of the distribution, such as the most probable top-item, the most probable rank-ing, or the distribution itself. More specifically, learning is done in an online manner, and the goal is to minimize sample complexity while guaran-teeing a certain level of confidence.

Preference-based reinforcement learning: evolutionary direct policy search using a preference-based racing algorithm [1]

We introduce a novel approach to preference-based reinforcement learn-ing, namely a preference-based variant of a direct policy search method based on evolutionary optimization. The core of our approach is a preference-based racing algorithm that selects the best among a given set of candidate policies with high probability. To this end, the algorithm operates on a suitable ordinal preference structure and only uses pairwise comparisons between sample rollouts of the policies. Embedding the racing algorithm in a rank-based evolutionary search procedure, we show that approxima-tions of the so-called Smith set of optimal policies can be produced with certain theoretical guarantees. Apart from a formal performance and complexity analysis, we present first experimental studies showing that our approach performs well in practice.

Biclique Coverings, Rectifier Networks and the Cost of ε-Removal [16]

We relate two complexity notions of bipartite graphs: the minimal weight biclique covering number Cov(G) and the minimal rec-tifier network size Rect(G) of a bipartite graph G. We show that there exist graphs with Cov(G) $\geq$ Rect(G) 3/2-ϙ . As a corollary, we estab-lish that there exist nondeterministic finite automata (NFAs) with ε-transitions, having n transitions total such that the smallest equivalent ε-free NFA has Ω(n 3/2-ϙ) transitions. We also formulate a version of previous bounds for the weighted set cover problem and discuss its con-nections to giving upper bounds for the possible blow-up.

Efficient Eigen-updating for Spectral Graph Clustering [2]

Partitioning a graph into groups of vertices such that those within each group are more densely connected than vertices assigned to different groups, known as graph clustering, is often used to gain insight into the organisation of large scale networks and for visualisation purposes. Whereas a large number of dedicated techniques have been recently proposed for static graphs, the design of on-line graph clustering methods tailored for evolving networks is a challenging problem, and much less documented in the literature. Motivated by the broad variety of applications concerned, ranging from the study of biological networks to the analysis of networks of scientific references through the exploration of communications networks such as the World Wide Web, it is the main purpose of this paper to introduce a novel, computationally efficient, approach to graph clustering in the evolutionary context. Namely, the method promoted in this article can be viewed as an incremental eigenvalue solution for the spectral clustering method described by Ng. et al. (2001). The incremental eigenvalue solution is a general technique for finding the approximate eigenvectors of a symmetric matrix given a change. As well as outlining the approach in detail, we present a theoretical bound on the quality of the approximate eigenvectors using perturbation theory. We then derive a novel spectral clustering algorithm called Incremental Approximate Spectral Clustering (IASC). The IASC algorithm is simple to implement and its efficacy is demonstrated on both synthetic and real datasets modelling the evolution of a HIV epidemic, a citation network and the purchase history graph of an e-commerce website.

From Bandits to Monte-Carlo Tree Search: The Optimistic Principle Applied to Optimization and Planning [36]

This work covers several aspects of the optimism in the face of uncertainty principle applied to large scale optimization problems under finite numerical budget. The initial motivation for the research reported here originated from the empirical success of the so-called Monte-Carlo Tree Search method popularized in computer-go and further extended to many other games as well as optimization and planning problems. Our objective is to contribute to the development of theoretical foundations of the field by characterizing the complexity of the underlying optimization problems and designing efficient algorithms with performance guarantees. The main idea presented here is that it is possible to decompose a complex decision making problem (such as an optimization problem in a large search space) into a sequence of elementary decisions, where each decision of the sequence is solved using a (stochastic) multi-armed bandit (simple mathematical model for decision making in stochastic environments). This so-called hierarchical bandit approach (where the reward observed by a bandit in the hierarchy is itself the return of another bandit at a deeper level) possesses the nice feature of starting the exploration by a quasi-uniform sampling of the space and then focusing progressively on the most promising area, at different scales, according to the evaluations observed so far, and eventually performing a local search around the global optima of the function. The performance of the method is assessed in terms of the optimality of the returned solution as a function of the number of function evaluations. Our main contribution to the field of function optimization is a class of hierarchical optimistic algorithms designed for general search spaces (such as metric spaces, trees, graphs, Euclidean spaces, ...) with different algorithmic instantiations depending on whether the evaluations are noisy or noiseless and whether some measure of the ”smoothness” of the function is known or unknown. The performance of the algorithms depend on the local behavior of the function around its global optima expressed in terms of the quantity of near-optimal states measured with some metric. If this local smoothness of the function is known then one can design very efficient optimization algorithms (with convergence rate independent of the space dimension), and when it is not known, we can build adaptive techniques that can, in some cases, perform almost as well as when it is known.

SERPICO Project-Team

6. New Results

6.1. Patch-based statistical denoising methods for electron and light microscopy

Participants: Charles Kervrann, Frédéric Lavancier.

Inspired form the non-local means [33], we developed a stochastic NL-means-based denoising algorithm for generalized non-parametric noise models [21], [9]. First, we provided a statistical interpretation to current patch-based neighborhood filters and justify the Bayesian inference that needs to explicitly account for discrepancies between the model and the data. Furthermore, we investigated the Approximate Bayesian Computation (ABC) rejection method [30], [46] combined with density learning techniques for handling situations where the posterior is intractable or too prohibitive to calculate. This is particularly relevant for images contaminated by heterogeneous sources of noise. A major difference with previous methods is that we directly handle the structure of the noise, without precise parametric modeling of the noise. We demonstrated the flexibility of our stochastic Gamma non-local means (SGNL-means) by showing how it can be adapted to tackle noise in frequency domain fluorescence lifetime imaging microscopy (FD-FLIM) and cryo-electron tomography (see Fig. 3).
Moreover, we also proposed a general statistical aggregation method which combines image patches denoised with several commonly-used algorithms [20]. We showed that weakly denoised versions of the input image obtained with standard methods, can serve to compute an efficient patch-based aggregated estimator. In our approach, we evaluate the Stein's Unbiased Risk Estimator (SURE) of each denoised candidate image patch and use this information to compute the exponential weighted aggregation (EWA) estimator. The aggregation method is flexible enough to combine any standard denoising algorithm and has an interpretation with Gibbs distribution. The denoising algorithm (PEWA) is based on an MCMC sampling and is able to produce results that are comparable to the state of the art ([4], [38]). In this range of work, we have also introduced in [28] a general method to combine estimators in order to produce a better estimate. From a theoretical point of view, we proved that this method is optimal in some sense. It is illustrated on standard statistical problems in parametric and semi-parametric models where the averaging estimator outperforms the initial estimators in most cases. As part of an on-going work, we are applying this method to improve patch-based image denoising algorithms.

References: [9] [21] [20] [28]

Collaborators: Philippe Roudot (UT Southwestern Medical Center, Dallas (TX))
Francois Waharte (UMR 144 CNRS-Institut Curie, STED team and PICT-IBiSA)
Paul Rochet (Laboratoire de Mathématiques Jean Leray (LMJL), university of Nantes)

6.2. Design of deconvolution algorithms for low exposure fluorescence microscopy images

Participants: Deepak George Skariah, Charles Kervrann.

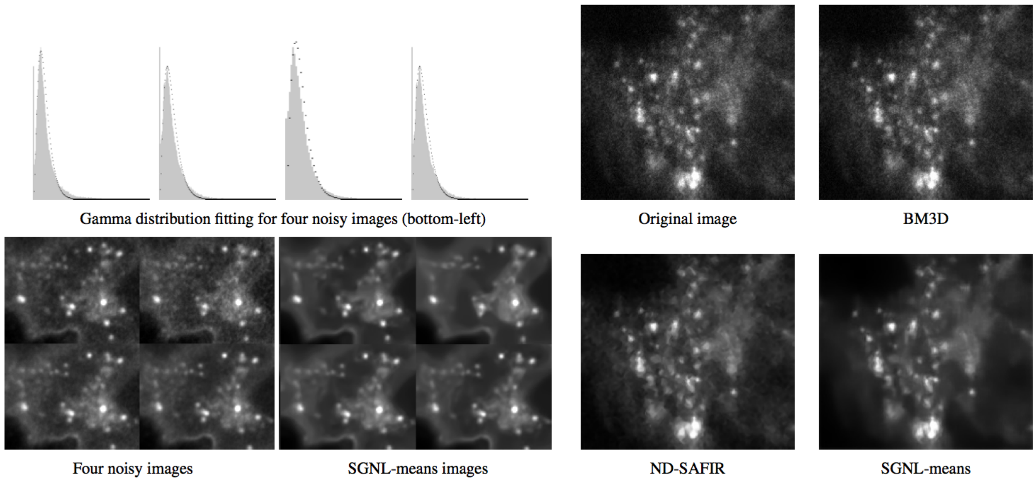


Figure 3. Experiments in FD-FLIM (confocal spinning-disk microscopy, UMR 144 CNRS-Institut Curie, PICT-IBiSA). Left: FNAR1 tagged with Green Fluorescence Protein (GFP) observed in a epithelial cell with mCHerry-tagged Tyk2 ; Gamma distribution fitting and SGNL-means denoising on four successive images with temporally varying signal-to-noise ratios. Right: comparison of denoised images with methods [4] [38].

Fluorescence imaging is popular in cell biology research due to its high contrast imaging capability. In microscopy imaging under low exposure conditions, the image quality is limited by out-of-focus blur and high noise. As a result a preprocessing stage known as deconvolution is needed to estimate a good quality version of the observed image. We proposed to design an efficient deconvolution algorithm for fluorescence microscopy under low exposure conditions by using the Poisson noise model. The result of deconvolution depends heavily on the choice of the regularization term. The regularization functional should be designed to remove noise while retaining the image structure. The choice of Poisson noise model and new regularization functional demands the design of a new and efficient optimization algorithm. We proposed to use a complex non quadratic regularization functional along with Poisson noise assumption for the first time. The use of non quadratic regularization makes the resulting optimization problem a complex one. This demanded the development of a problem-specific optimization algorithm which is fast as well as robust enough to minimize a non quadratic cost function. The use of non quadratic regularization together with Poisson noise model ensures that finer details of underlying structures are well restored in the presence of large amount of noise.

Collaborator: Muthuvel Arigovindan (Imaging Systems Lab, Department of Electrical Engineering, Indian Institute of Science, Bangalore, India).

6.3. Background estimation and vesicle segmentation in live cell imaging

Participants: Thierry Pécot, Patrick Bouthemy, Charles Kervrann.

In live cell fluorescence microscopy images, the moving tagged structures of interest, such as vesicles, often appear as bright spots with intensity that varies along time over a time-varying and cluttered background. Localization and morphology assessment of these small objects over time is then crucial to provide valuable information for quantitative traffic analysis. In this study, we have focused on the Rab6 protein as a typical intracellular membrane-associated protein. Rab6 is known to promote vesicle trafficking from Golgi to Endoplasmic Reticulum or to plasma membrane. In our study, micro-fabricated patterns have been used to enforce cells to have circular or crossbow normalized shape. Micro-patterns impose constraints on the cytoskeleton and the location of organelles (e.g. Golgi apparatus) is thus better controlled. These micro-patterns also influence the spatial distribution of Rab6 transport carriers. However, the direct influence of the micro-patterns on the spatial dissemination of these trafficking vesicles has so far not been completely characterized. In this work, we have considered a statistical Bayesian approach in the framework of conditional random fields (CRF) for background estimation and vesicle segmentation [13]. Within this approach, we have designed a robust detection measure for fluorescence microscopy based on the distribution of neighbor patch similarity. We formulate the vesicle segmentation and background estimation as a global energy minimization problem. An iterative scheme to jointly segment vesicles and background is proposed for 2D-3D fluorescence image sequences. We have conducted a quantitative comparison with state-of-the-art methods on a large set of synthetic image sequences with a cluttered time-varying background and achieved a quantitative validation of the vesicle segmentation method on 2D and 3D micro-patterned cells expressing GFP-Rab6.

Reference: [13]

Collaborators: Jean Salamero (UMR 144 CNRS-Institut Curie, STED team and PICT-IBiSA)
Jérôme Boulanger (UMR 144 CNRS-Institut Curie, STED team)

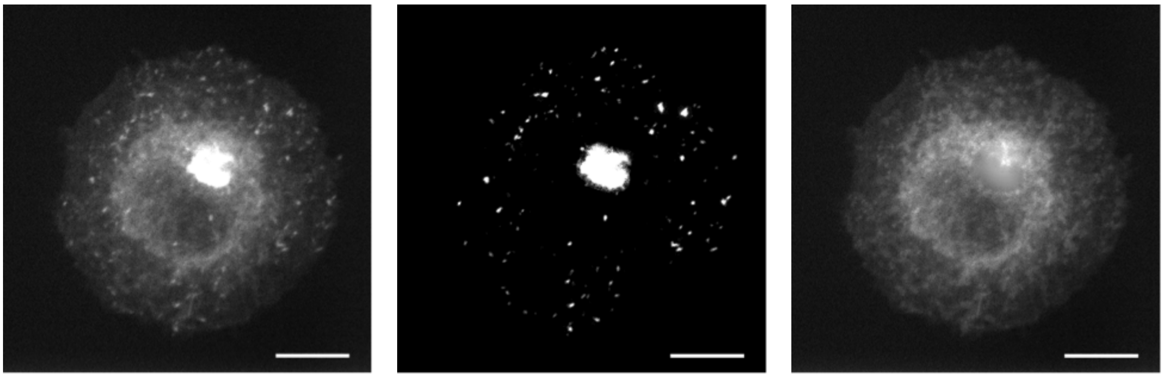


Figure 4. Left: Fuorescence confocal spinning-disk microscopy image depicting GFP-Rab6 proteins (UMR 144 CNRS-Institut Curie, PICT-IBiSA). Middle: estimated vesicular component. Right: estimated background.

6.4. A quantitative approach for space-time membrane trafficking orientation

Participants: Thierry Pécot, Patrick Bouthemy, Charles Kervrann.

Rab6 proteins are trafficking from the Golgi apparatus at the cell center to Endoplasmic Reticulum or to plasma membrane located at the periphery of the cell. The cell shape influences Rab6 trafficking but no study has ever quantified the effect of the cell shape on the trafficking orientation. In this study [25], we compare Rab6 trafficking orientation constrained by two different micropatterns [56] (circular and crossbow-shaped cells) from fluorescence video-microscopy. Object/background separation [13] is first applied to 3D+T image sequences to extract Rab6 spatio-temporal coordinates. The bandwidth of the von Mises kernel is automatically estimated using the rule of thumb and leads to two different densities for the two different micropatterns. We propose to quantitatively compare these densities by computing the Wilcoxon rank sum paired test between inter- and intra-micropattern distances. We considered the circular earth mover's distance (also known as the Wasserstein metric) to compare traffic densities. Our quantitative study on micro-patterned cells concludes that the Rab6 transport carriers destinations concentrate at the three corner points of the crossbow-shaped cells corresponding to the main adhesion sites, while the vesicle destination distribution is somewhat uniform for circular-shaped cells.

Reference: [25]

Collaborators: Jean Salamero (UMR 144 CNRS-Institut Curie, STED team and PICT-IBiSA)
Jérôme Boulanger (UMR 144 CNRS-Institut Curie, STED team)

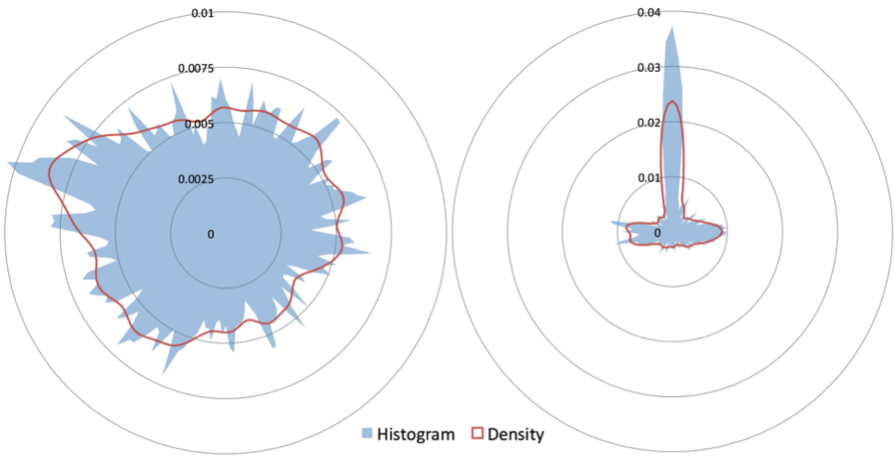


Figure 5. Distribution of traffic orientation for circle-shaped cells (left) and crossbow-shaped cells (right).

6.5. Vesicle segmentation method with automatic scale selection in TIRF microscopy

Participants: Antoine Basset, Charles Kervrann, Patrick Bouthemy.

Accurately detecting subcellular particles in fluorescence microscopy is of primary interest for further quantitative analyses such as counting, tracking or classification. Our primary goal was to segment vesicles in fluorescence microscopy images. In [15] we proposed a first spot detection method with automatic scale selection. We have now dramatically improved the precision of the scale selection step, yielding to a more reliable detection of the spots [23]. The method relies on a Laplacian of Gaussian (LoG) filter to first enhance the spots while reducing noise. To obtain good detection results, the scale of the Gaussian filter must be precisely set, according to the spots size [23]. In order to cope with very small spots, we rely on the discrete analog of the Gaussian filter [45], instead of the previously used sampled Gaussian filter. With this filter, we can find the optimal Gaussian scale with an arbitrary precision by minimizing a statistical criterion. We have introduced two criteria for this purpose and compared them. Once the optimal scale is selected, we threshold the lowest values of the LoG-filtered image, which correspond to spots. To cope with inhomogeneous background, thresholding must be adapted to local statistics so that a single probability of false alarm (PFA) setting can be defined for the whole image or even the collection of images to be processed. In short, we automatically infer from image data the optimal parameters usually left to the user guidance in other methods, that is, spot scale and detection threshold. We have carried out an extensive comparative evaluation, which demonstrates that our new scale selection approach improves detection performances, and that our spot detection method outperforms state-of-the-art detectors [23].

References: [15] [23]

Collaborators: Jean Salamero (UMR 144 CNRS-Institut Curie, STED team and PICT-IBiSA)
Jérôme Boulanger (UMR 144 CNRS-Institut Curie, STED team)

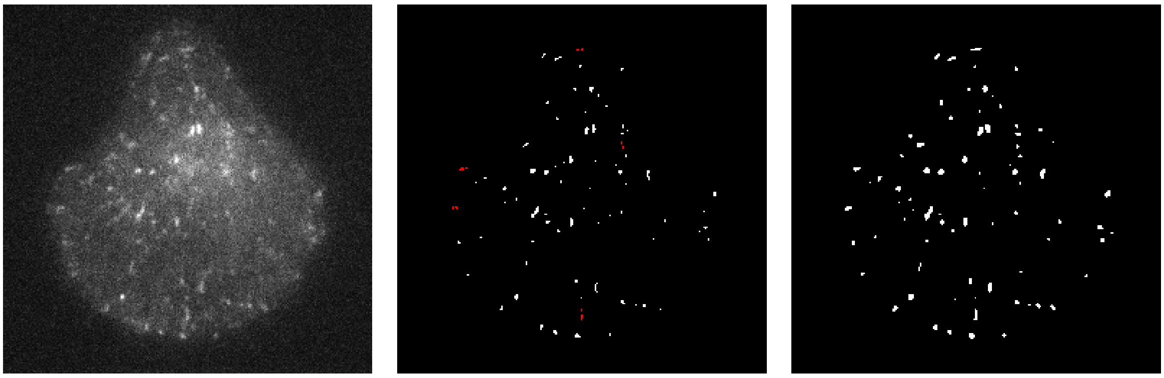


Figure 6. Comparison of segmentation results on a real image presenting elongated spots. Left: Input TIRFM images (Rab11-mCherry) (UMR 144 CNRS-Institut Curie, PICT-IBiSA). Middle: Segmentation results with state-of-the-art detector MS-VST [60]. Some elongated spots of (left) are split (red) by MS-VST due to a too small filter scale. Right: Segmentation results with our new detection method. Elongated objects are well recovered thanks to the precise scale selection.

6.6. Analysis of the repartition of moving vesicles by spatio-temporal point process models

Participants: Frédéric Lavancier, Thierry Pécot, Charles Kervrann.

Characterizing the spatial repartition of interacting moving proteins is a fundamental step for co-localization and co-expression. Based on the segmentation algorithm [15], [23], this challenge amounts to characterizing the repartition or spatial distribution of spots (see Fig. 6). This is part of the more general statistical analysis of random geometrical objects, and in particular of random points. Gibbs models form a large class of point process models, that can be used to characterize either complete randomness or attraction or repulsion between points depending on the Gibbs potential at hand.

First in [27], we focused on infinite range potentials that include the most famous interaction potential arising from statistical physics, namely the Lennard Jones potential. To fit this kind of models to a dataset, the standard inference methods are not applicable. We introduced in [27] a modification of the pseudolikelihood method, with a specific border correction, and we prove that this provides consistent and asymptotically normal estimators. Second, in [26], we studied an alternative class of models, the determinantal point processes (DPP). They are designed to model repulsion between points and are thus adapted to regular point patterns. These models are becoming very popular in the spatial statistics community due to many appealing properties. We quantified the possible repulsiveness that a DPP can model [26]. In particular, we determined the most repulsive stationary DPP. We finally introduced new parametric families of DPPs that cover a large range of DPPs, from the homogeneous Poisson process (for no interaction) to the most repulsive DPP.

An application of these models to the problem of co-localization between proteins is part of an on-going project. In each protein, the set of vesicles is modeled by a union of random balls, possibly overlapping, and a Gibbs interaction is introduced to take into account the possible interaction in the location of vesicles between two proteins. Our first concern is to test whether the two proteins actually interact, i.e. co-localization occurs, or in other words whether the Gibbs interaction is empty or not. If there is co-localization, the further step is to characterize it through the estimation of the strength of the Gibbs interaction.

References: [26] [27]

Collaborators: Christophe Ange Napoléon Biscio (LMJL, University of Nantes)

Jean-François Coeurjolly (Laboratoire Jean Kuntzmann, Grenoble Alpes University)

6.7. Detection and estimation of membrane diffusion during exocytosis in TIRF microscopy

Participants: Antoine Basset, Charles Kervrann, Patrick Bouthemy.

Assessing the dynamics of plasma membrane diffusion processes in live cell fluorescence microscopy is of paramount interest to understand cell mechanisms. We investigated methods to detect vesicle fusion events, and estimate the associated diffusion coefficients in TIRFM image sequences [16]. In contrast to classical approaches, a diffusion coefficient is locally estimated for each detected fusing vesicle. We first detect the membrane fusion events and then select the diffusion configurations among them with a correlation test. To estimate the diffusion coefficient, a geometric model is fitted to the detected spot directly in the 2D+T subvolume. This recent estimation approach produced more satisfying results when compared to [16]. Diffusion events are reliably recognized, and the diffusion coefficient is accurately estimated for each diffusion event. This work will be integrated in a broader study, spanning from transport phase to membrane fusion, and non-diffusion events will be analyzed.

Reference: [16]

Collaborators: Jean Salamero (UMR 144 CNRS-Institut Curie, STED team and PICT-IBiSA)
Jérôme Boulanger (UMR 144 CNRS-Institut Curie, STED team)

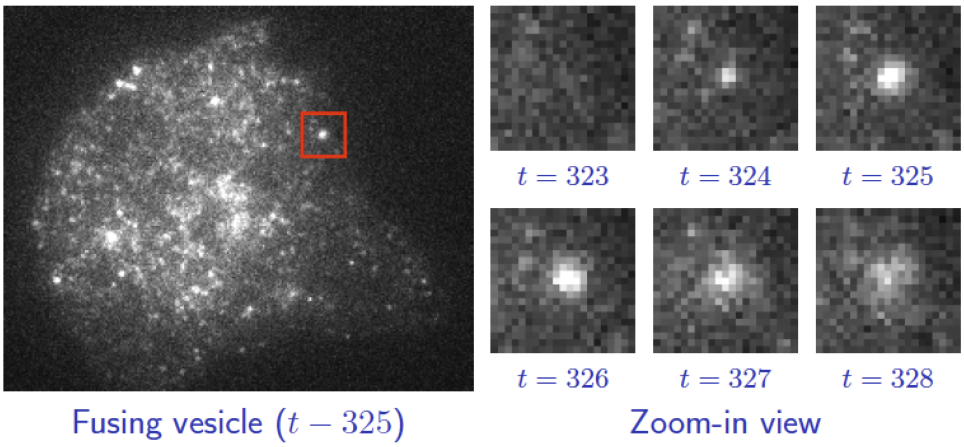


Figure 7. Left: Fusing vesicle (frame in red) in a TIRFM (UMR 144 CNRS-Institut Curie, PICT-IBiSA) sequence (frame 325, 50ms/frame). Right: Zoom-in view of the temporal evolution of the fusing vesicle.

6.8. Estimation of the flow of particles without tracking in fluorescence video-microscopy

Participants: Thierry Pécot, Patrick Bouthemy, Charles Kervrann.

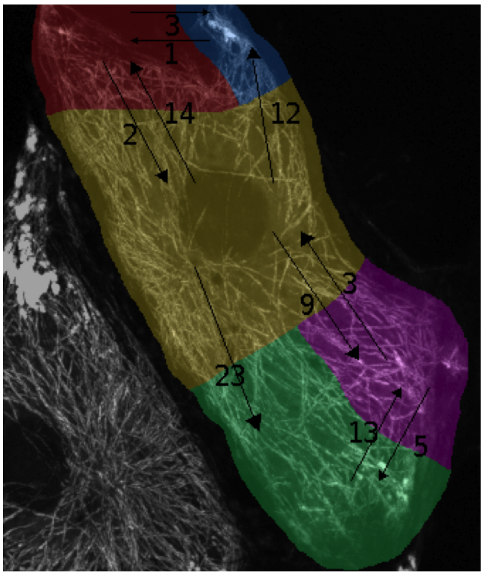


Figure 8. Vesicle flows estimated when considering a simple partition of 5 regions for an image sequence acquired in TIRF microscopy and showing the protein Clip170 (UMR 144 CNRS-Institut Curie, PICT-IBiSA).

Automatic analysis of the dynamic content in fluorescence video-microscopy is crucial for understanding molecular mechanisms involved in cell functions. We have proposed an original approach for analyzing particle trafficking in these sequences. Instead of individually tracking every particle, we only locally count particles crossing boarders between regions over time and minimize a global energy function. Three methods to determine the particle flow have been considered. We have conducted comparative experiments on synthetic and real fluorescence image sequences. We have shown that adding a sparsity constraint on the number of detected events allows us to reduce the number of false alarms. Compared to usual tracking methods, our approach is simpler and the results are very stable. This estimation method needs the adjustment of only two parameters. (see Fig. 8).

Reference: [22]

Collaborators: Jean Salamero (UMR 144 CNRS-Institut Curie, STED team and PICT-IBiSA)
Jérôme Boulanger (UMR 144 CNRS-Institut Curie, STED team)

6.9. Detection and tracking of astral microtubules at the cell cortex

Participants: Thierry Pécot, Charles Kervrann, Geoffrey Dieffenbach.

In this study, we are interested in the influence of the mechanical properties of astral microtubules in the centering mechanisms of the mitotic spindle, giving it a robust positioning. In their previous studies, the CeDRE group (IGDR Rennes) identified two subpopulations of astral microtubules that either push or pull the cell cortex. To better understand these mechanisms, image sequences are acquired at the cortex level where extremities of astral microtubules come to exert forces. In order to characterize the two subpopulations of astral microtubules during the mitosis in the unicellular embryos of C. Elegans, life span, that is the period during which the microtubule is touching the cell cortex, for every single microtubule has to be measured. A short life span corresponds to a pulling force while a longer life span corresponds to a pushing force. Detecting and tracking microtubules at the cell cortex has to be done to collect these measures. As the signal-to-noise ratio is low, a denoising step is needed to detect the microtubule extremities. Several detection methods were tested but we need to further investigate this step to find the most suited methods for this particular application. Finally, the U-track algorithm [42] is applied to track the microtubules extremities to measure their life span.

Collaborators: Jacques Pécréaux (CeDRE group, IGDR Rennes, CNRS UMR 6290)
Hélène Bouvrais (CeDRE group, IGDR Rennes, CNRS UMR 6290)

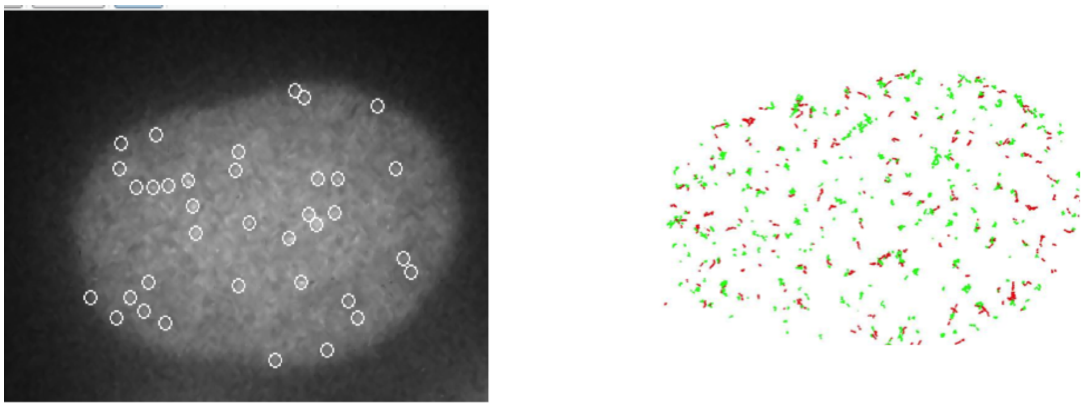


Figure 9.

Microtubule extremities detection and tracking in fluorescence microscopy (embryo of C. Elegans, IGDR - Institute of Genetics and Developmental biology of Rennes, CNRS UMR 6290).

6.10. Spot localization and segmentation for Tissue MicroArray (TMA) de-arrying

Participants: Hoai Nam Nguyen, Charles Kervrann.

Tissue core de-arraying is one of the most important steps in tissue microarray (TMA) image analysis. A very first task of TMA (Tissue MicroArray) image analysis is to accurately localize spots (separate tissue core) representing arrays of 512×512 pixels each, in very large images of several thousands of pixels. However, few solutions and frameworks are available and none of them covers images provided by fluorescent scanners. We developed a robust TMA de-arraying method adapted for digital images from classical optical and new fluorescent devices. The proposed algorithm is composed of three modules: i) detection, ii) segmentation, and iii) array indexing. The detection of TMA cores is performed by local adaptive thresholding of isotropic wavelet transform coefficients. We demonstrated how a wavelet decomposition at any desired scale can be performed faster than usual techniques by exploiting explicit formula of the analysis wavelet. Our core detection strategy enables to deal with images having significant noise level, inhomogeneous background, and high dynamic range such as fluorescence images, without any assumption on image noise and intensity value range. The detected cores are furhter segmented by using parametric ellipse model to improve detection accuracy. Combining these two modules, we can handle complex background and artifacts, particularly in fluorescence imaging, and thus reduce false detections. After the segmentation step, the position of detected cores is determined by the centroid of relevant segments. Finally, to compute array indices of cores, we estimate the deformation of a theoretical grid under a thin-plate model by using an iterative scheme. After each iteration, the initial regular grid is progressively transformed for fitting computed core positions. Our main contribution is the reformulation of the array indexing problem as an estimation of the deformation function, which is solved with a iterative algorithm. Moreover, when design layout of TMA slide is known, our estimator of deformation yields quantitative information about grid deformation such as average translation, rotation angle, shearing coefficients, bending energies along axis, etc. They can be used as quality indicators of the manufactured TMA slide.

Collaborator: Vincent Paveau (Innopys company)

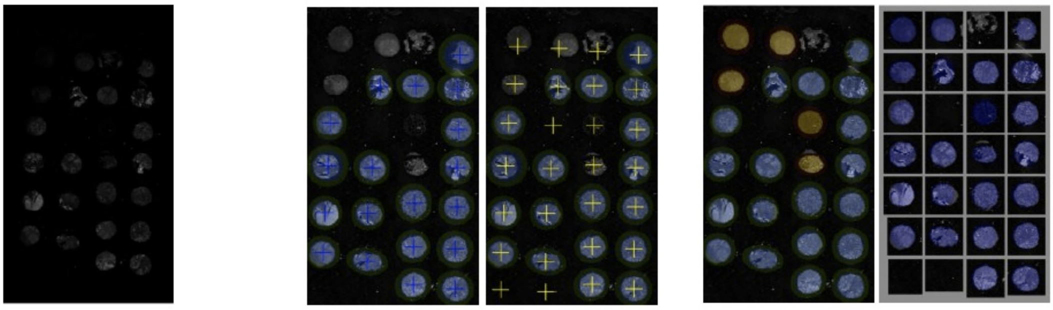


Figure 10. Array indexing TMA (Innopsys company). From left to right : input TMA image, segmented core positions marked by blue crosses, estimated positions of deformed grid marked by yellow crosses, retrieved missed cores after detection/segmentation steps (orange areas), and array representation of TMA (retrieved cores are colored).

6.11. Adaptive global and local motion estimation

Participants: Noémie Debroux, Charles Kervrann.

The design of data costs is one of the main research issue for variational optical flow estimation. The aim is to improve discriminative power by integrating appropriate neighborhood information, while preserving computational efficiency. Most previous works define features on patches with predefined sizes and shapes, or filter pixelwise costs with fixed filtering parameters. We proposed a novel approach estimating spatially varying parameters of filters used to define the data term [8]. More specifically, our model considers Gaussian filtering of the pixelwise brightness constancy equation and imposes smoothness constraints on motion and convolution filter size (bandwith). The energy encoding these assumptions is alternatively minimized over flow field and the spatially varying bandwidth in a variational framework. Experimental results on the Middlebury database demonstrated clear improvements yielded by our method over the spatially constant case of [32] (see Fig. 11).

Collaborator: Denis Fortun (UMR 144 CNRS-Institut Curie, STED team, Paris)
(EPFL, Lausanne, Switzerland)

6.12. Crowd motion classification

Participants: Antoine Basset, Charles Kervrann, Patrick Bouthemy.

Assessing crowd behaviors from videos is a difficult task while of interest in many applications. We have defined a novel approach which identifies from two successive frames only, crowd behaviors expressed by simple image motion patterns. It relies on the estimation of a collection of sub-affine motion models in the image, a local motion classification based on a penalized likelihood criterion, and a regularization stage involving inhibition and reinforcement factors [17]. The apparent motion in the image of a group of people is assumed to be locally represented by one of the three following motion types: translation, scaling or rotation. The three motion models are computed in a collection of predefined windows with the robust estimation method [48]. At every point, the right motion model is selected owing to the corrected (for small sample size) Akaike information criterion (AICc). To classify the local motion type, the three motion models are further subdivided into a total of eight crowd motion classes. Indeed, scaling refers either to gathering (Convergence) or dispersing people (Divergence). Rotation can be either Clockwise or Counterclockwise.

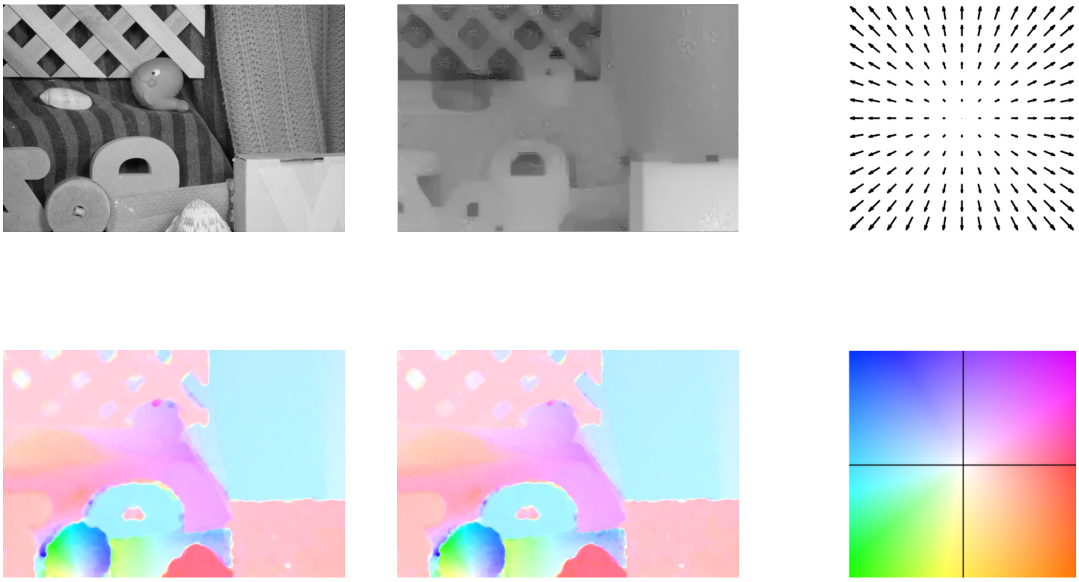


Figure 11. Comparison on a sequence of the Middlebury benchmark. Top from left to righ: input image and spatially filter bandwidth estimation. Bottom from left to right: velocity field computed by [32] (endpoint error = 0.143) and by our method (endpoint error = 0.126).

Since our classification scheme is view-based, four image-related translation directions are distinguished: North, West, South, East. Then, to get the final crowd classification, a regularization step is performed, based on a decision tree and involving inhibition for opposed classes such as convergence and divergence. We have also developed an original and simple method for recovering the dominant paths followed by people in the observed scene. It involves the introduction of local paths determined from the space-time average of the parametric motion subfields selected in image blocks. Starting from one given block in the image, we straightforwardly reconstruct a global path by concatenating the local paths from block to block. Experiments on synthetic and real scenes have demonstrated the performance of our method, both for motion classification and principal paths recovery.

Reference: [17]

6.13. Anomaly detection using block-based histograms of crowd motion patterns

Participants: Juan Perez Rua, Antoine Basset, Patrick Bouthemy.

We have developed a new and generic method to detect and localize abnormal events in videos of crowd scenes. The algorithm consists first in determining the flow vector and crowd motion class for every moving pixel from a set of affine motion models estimated on a collection of windows. Then, the observed scene is subdivided in blocks to compute crowd motion class histograms weighted by the motion vector magnitudes. A very simple training step enables to get the reference histograms per block accounting for the normal behaviours. For each block, we can automatically set by means of statistical arguments the threshold on the distance between the histogram in the current image and the reference histogram that decides the presence of an abnormal event in that block. Results of extensive experimentation on different types of anomaly datasets show that our method is competitive with respect to methods relying on far more elaborated models on both appearance and motion and thus involving a significant learning stage. It outperforms any other existing purely motion-based anomaly localization method.

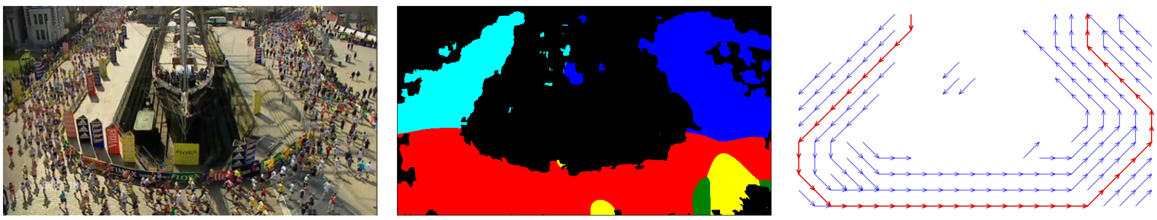


Figure 12. Overview of the method applied to a sequence where runners follow a ‘U’ from the upper left corner to the upper right corner. Left: First frame of the sequence. Middle: Classification results (cyan=translation toward South, red=counterclockwise rotation, yellow=translation toward East, green=convergence, blue=translation toward North). Right: Recovery of the longest path in the scene (red).

SHACRA Project-Team

5. New Results

5.1. Highlights of the Year

5.1.1. Intra-operative guidance

Each year in Europe 50,000 new liver cancer cases are diagnosed for which hepatic surgery combined to chemotherapy is the most common treatment. In particular the number of laparoscopic liver surgeries has increased significantly over the past years. Minimally invasive procedures are challenging for the surgeons due to the limited field of view.

Providing new solutions to assist surgeons during the procedure is of primary interest. This year, the team developed an innovative system for augmented reality in the scope of minimally invasive hepatic surgery. The first issue is to align preoperative data with the intra-operative images. We first proposed a semi-automatic approach [28] for solving the ill-posed problem of initial alignment for augmented reality systems during liver surgery. Our registration method relies on anatomical landmarks extracted from both the laparoscopic images and a three-dimensional model, using an image-based soft-tissue reconstruction technique and an atlas-based approach, respectively.

Second, we introduced a method for tracking the internal structures of the liver during robot-assisted procedures [25]. Vascular network, tumors and cut planes, computed from pre-operative data, can be overlaid onto the laparoscopic view for image-guidance, even in the case of large motion or deformation of the organ. This is made possible by relying on a fast yet accurate 3D biomechanical model of the liver combined with a robust visual tracking approach designed to properly constrain the model. Our augmented reality proved to be accurate and extremely promising on in-vivo sequences of a human liver during robotic surgery.

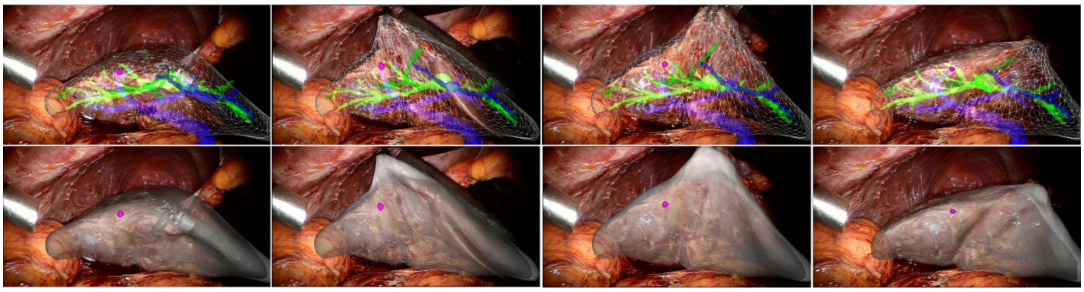


Figure 4. Augmented reality on the liver with 3D visualization of the blood vessels

5.1.2. Ph.D. defenses

The year 2014 was also special since many PhDs have been defended. Four PhD defenses took place with:

- Ahmed Yureidini's defense about *Robust blood vessel surface reconstruction for interactive simulations from patient data* [15] in May 2014,
- Guillaume Kazmitcheff's defense about *Minimal invasive robotics dedicated to otological surgery* [13] in June 2014,
- Hugo Talbot's defense about *Interactive patient-specific simulation of cardiac electrophysiology* [14] in July 2014,
- Alexandre Bilger's defense about *Patient-specific biomechanical simulation for deep brain stimulation* [12] in December 2014.

5.1.3. *Organization of ISBMS 2014*

The team co-organized the 6^{th} International Symposium on Biomedical Simulation (ISBMS) 2014, which was held in Strasbourg (France) on October 16 – 17, 2014. The ISBMS conference is a well-established scientific meeting that provides an international forum for researchers interested in using biomedical simulation technology for the improvement of patient care and patient safety. The SiMMS group from Imperial College London and IHU-Strasbourg were the two other co-organizers. The event was hosted at IRCAD, a center of excellence in surgical training. The ISBMS chairs were:

- Stéphane Cotin (Inria),
- Fernando Bello (Imperial College London),
- Jérémie Dequidt (Univ. Lille),
- Igor Peterlik (IHU Strasbourg & Masaryk Univ.).

The whole team was involved in the organization of the event. About 65 participants joined the conference. Regarding their feedback, the conference was a real success. For more information about ISBMS, refer to the official website http://www.isbms.org.

Finally, a day dedicated to our software SOFA ("SOFA Day") was organized the day after the ISBMS conference. This was the opportunity to introduce SOFA to the ISBMS community and to share with the SOFA users.

(a) Setup of our demo (b) With Genevieve Fioraso

Figure 5. Presentation of our work at the French National Assembly. Genevieve Fioraso is the French national research secretary

5.1.4. *Demonstration at the French National Assembly*

On Tuesday 21^{st} January 2014, the team SHACRA presented its work during the "Internet et société numérique" working group. This was a joint event between Inria and the French National Assembly (Assemblée Nationale). On this special occasion, we made a demonstration of our simulations and the CEO from Inria Michel Cosnard also presented more globally the role of Inria in healthcare but also education, cloud computing, big data.

5.2. New Results

5.2.1. *Real-Time Biophysical Models*

5.2.1.1. Deep brain stimulation

Participant: Alexandre Bilger.

During this year, we developed an intra-operative registration method. It is used during a DBS surgery and can help the surgeon to locate anatomical structures for a safer and a more efficient treatment [21]. The method is based on the biomechanical model of brain shift we developed during the last years. Because some parameters of the model are unknown, we propose to estimate them with an optimization process. The cost function evaluates the distance between the model and the segmentation of pneumocephalus, the only indicator of brain shift visible on an intra-operative CT scan.

Figure 6. Biomechanical model of the brain for DBS planning

5.2.1.2. Stapedectomy

Participant: Guillaume Kazmitcheff.

Stapedectomy is a challenging procedure of the middle ear microsurgery, since the surgeon is in direct contact with sensitive structures such as the ossicular chain. This procedures is taught and performed in the last phase of the surgical apprenticeship. To improve surgical teaching, we propose to use a virtual surgical simulator [26] based on a finite element model of the middle ear. The static and dynamic behavior of the developed finite element model was successfully compared to published data on human temporal bones specimens. A semi-automatic algorithm was developed to perform a quick and accurate registration of our validated mechanical atlas to match the patient dataset. This method avoids a time-consuming work of manual segmentation, parameterization, and evaluation. A registration is obtained in less than 260 seconds with an accuracy close to a manual process and within the imagery resolution. The computation algorithms, allowing carving, deformation of soft and hard tissus, and collision response, are compatible with a real-time interactive simulation of a middle ear procedure. As a future work, we propose to investigate new robotized procedures of the middle ear surgery in order to develop new applications for the RobOtol device and to provide a training tool for the surgeons.

5.2.1.3. Cardiac electrophysiology

Participant: Hugo Talbot.

Cardiac arrhythmia is a very frequent pathology that comes from an abnormal electrical activity in the myocardium. The skills required for such interventions are still very challenging to learn, and typically acquired over several years. We first developed a training simulator for interventional electrocardiology and thermo-ablation of these arrhythmias [14], [32]. Based on physical models, this training system reproduces the different steps of the procedure, including endovascular navigation, electrophysiological mapping, pacing and cardiac ablation. Based on a scenario of cardiac arrhythmia, cardiologists assessed the interactivity and the realism of our simulation.

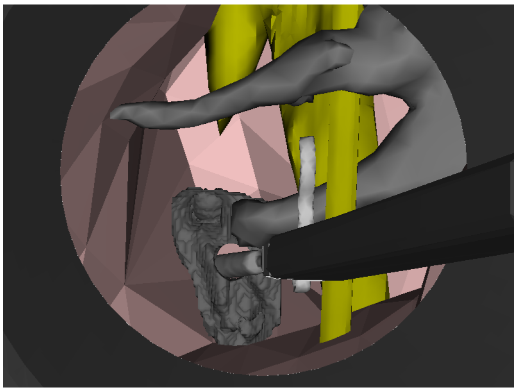


Figure 7. Simulation of the stapedotomy procedure

Figure 8. Training simulator for electrocardiology procedures

Beyond electrophysiology training, our work around the cardiac electrophysiology also target the personalization of our mathematical models. Using the dense electrograms recorded intra-operatively, we presented an accurate and innovative approach to personalize our model, i.e. estimate patient-specific parameters. The modeling in silico of a patient electrophysiology is needed to better understand the mechanism of cardiac arrhythmia. This work has been submitted in a conference.

5.2.1.4. Cryoablation

Participant: Hugo Talbot.

A new project started this year around cryotherapy. This technique consists in inserting needles that freezing the surrounding tissues, thus immediately leading to cellular death of the tissues. Cryoablation procedure is used in many medical fields for tumor ablation, and even starts being used in cardiology. In this scope, we build a simulator able to place the cryoprobes and run a simulation representing the evolution of iceballs in living tissues [31]. This work was presented at MMVR'14.

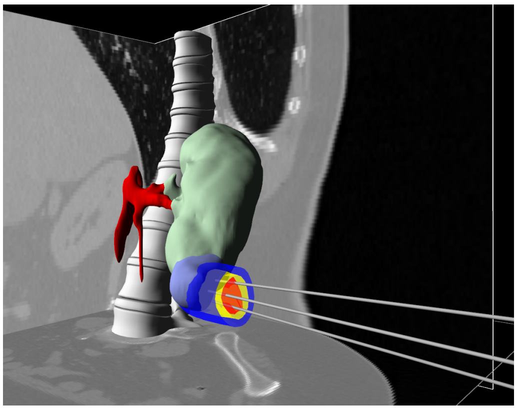


Figure 9. Simulation of the stapedotomy procedure

5.2.1.5. Connective tissues

Participant: Julien Bosman.

Another topic of simulation is the modeling of connective tissues [18]. First, a comparative study on the influence of the ligaments in liver surgery has been conducted. This study underlines that the model chosen for the ligament's has a strong influence on the outcome of the simulation. More specifically, it shows the the model is at least as much important as the material parameters of the parenchyma. It also shows that the influence of the model depends on the type of effort that is prescribed on the liver. The second axis concerns the validation of a frame (6-DOF nodes) based mechanical model developed for ligaments simulation. Current results show that this model requires less degrees of freedom while providing the same accuracy as a traditional FEM model. At last, a method dedicated to the simulation and the control of continuum robots has been developed. The goal of this method is to replace the mesh of robot by computing its compliance and applying it on a reduced model made of frames. It allows to strongly decrease the number of degrees of freedom needed for the robot simulation while keeping the needed accuracy.

5.2.1.6. Simulation of lipofilling reconstructive surgery

Participant: Vincent Majorczyk.

We have developed a method to simulate the outcome of reconstructive facial surgery based on fat-filling. Facial anatomy is complex: the fat is constrained between layers of tissues which behave as walls along the face; in addition, connective tissues that are present between these different layers also influence the fat-filling procedure. To simulate the end result, we have proposed a method which couples a 2.5D Eulerian fluid model for the fat and a finite element model for the soft tissues. The two models are coupled using the computation of the mechanical compliance matrix. We had two contributions: a solver for fluids which couples properties of solid tissues and fluid pressure, and an application of this solver to fat-filling surgery procedure simulation.

5.2.1.7. Inverse FEM simulation

Participant: Eulalie Coevoet.

We introduced a new methodology for semi-automatic deformable registration of anatomical structures [23], using interactive inverse simulations. We applied the approach for the registration of the parotid glands during the radiotherapy of the head and neck cancer. Radiotherapy treatment induces weight loss that modifies the shape and the positions of these structures and they eventually intersects the target volume. We proposed a method to adapt the planning to limit the radiation of these glands.

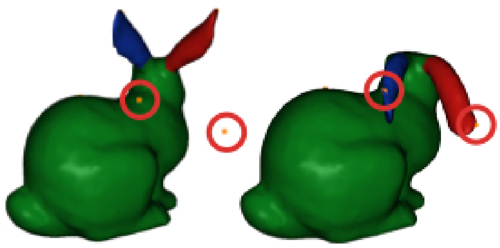


Figure 10. Numerical validation. Left: Target points (highlighted in red) after setting 3 different Young's moduli (one color by Young's modulus). Right: The resulting deformation once the Young modulus have been estimated.

5.2.2. Numerical Methods for Complex Interactions

5.2.2.1. Cliping in neurosurgery

Participant: Eulalie Coevoet.

We developed a simulator for neurosurgery. The surgery consist in "clipping" a cerebral aneurysm. Aneurysm is an abnormal local dilatation in the wall of a blood vessel, usually an artery. There are several treatment options for people with the diagnosis of cerebral aneurysm. Medical therapy, surgical therapy (clipping) and endovascular therapy (coiling). The surgical therapy, because of his invasive and technical nature, is the less prescribed. This leads to less and less surgeon trained to practice the procedure. And yet some patients require the surgical way. So the idea was to develop a simulator to train student and also help on the planification.

5.2.2.2. Virtual cutting

Participants: Huu Phuoc Bui, Christoph Paulus.

The simulation of cutting is a central interest in the team. Several approaches have been investigated this year to model surgical cuts, tearing and other separations of materials induced by surgical tools:

- using the standard finite element method (FEM) combined with a re-meshing approach, that replaces locally the current structure of the mesh in order to allow for a separation,
- using the extended FEM (X-FEM) that uses shape functions that can model discontinuities inside elements (see Fig. 12),
- and using the Lattice element method (LEM).

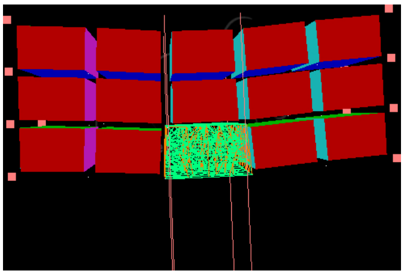
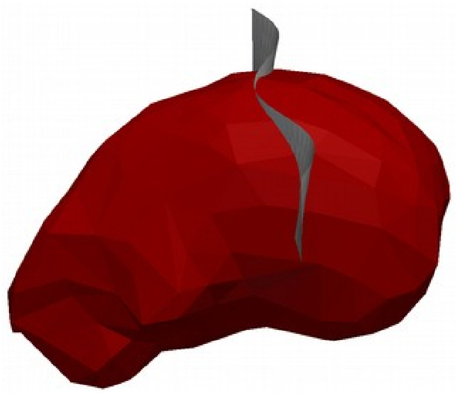


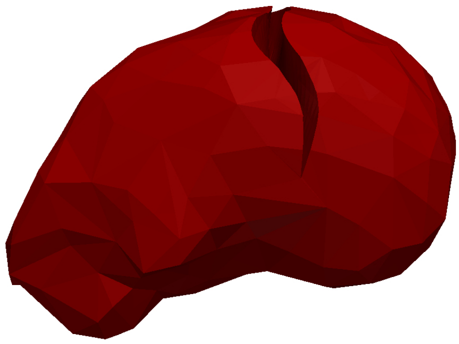
Figure 11. Cutting simulation using LEM

A re-meshing approach to model cuts has been submitted to several conferences, we are waiting for the response. An implementation of the extended finite element method was published in a preprint "Simulation of Complex Cuts in Soft Tissue with the Extended Finite Element Method (X-FEM)". The figures below show a simulation of a sinusoidal cut on a liver executed with the implementation of the X-FEM.

For the LEM approach (see Fig. 11), a multimapping between finite elements and lattice model have been developed and implemented into SOFA framework. This allows us to perform a multiscale simulation in real-time. A dynamic changing of topology between finite elements and lattices should be developed in the next step in order to perform the cutting dynamically.



(a) Plane of cut



(b) Result of the cutting

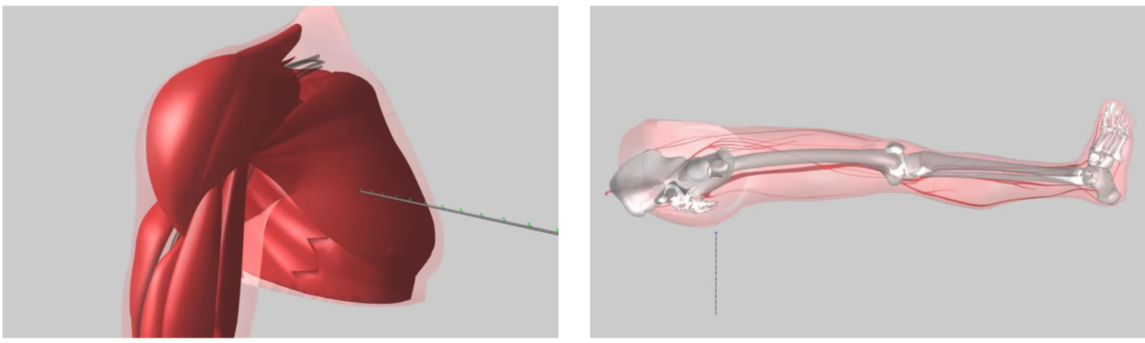
Figure 12. Cutting simulation based on X-FEM

5.2.2.3. Regional anaesthesia

Participants: Rémi Bessard Duparc, Frédérick Roy.

The RASimAs project (Regional Anaesthesia Simulator and Assistant) is a European research project funded by the European Union's 7th Framework Program. It aims at providing a virtual reality simulator and assistant to doctors performing regional anaesthesia by developing the patient-specific Virtual Physiological Human models. This year, the code for needle insertion has been re-designed and simplified into SOFA and the muscle contraction has been implemented. Finally, the components of the simulation have been optimized to reach the desired real-time performances (i.e more than 25-30 frames per second).

Our preliminary results are awaiting the validation of the Working Packages in January 2015. The needle refactoring will be shared with an other project in Strasbourg (robot) and may be shared with an other team at Inria Rennes with the LAGADIC Team.



(a) Needle insertion in the shoulder (b) Needle insertion in the leg

Figure 13. Regional anaesthesia with needle insertion and muscle contraction

5.2.2.4. Control of elastic soft robots

Participant: Frédérick Largillière.

We developed a prototype of stiffness-controlled haptic interface using a piece of silicone rubber to render different forces related to a displacement ie. different stiffnesses and an improved method of simulation using multi-rate loops to try to keep the computation real-time even with models using a large number of FEM elements. (work currently under review) We also presented the idea of a surgical robot able to virtually reconstruct its environment (ie. surrounding biological tissues) through small modifications of the algorithm used for controlling soft robots (SURGETICA 2014).

5.2.3. Image-Driven Simulation

5.2.3.1. Physics-based registration algorithms

Participant: Rosalie Plantefève.

Before targeting the augmented reality for laparoscopic operations, an important step consists in solving the initial alignment problem. Given a pre-operative image of the organ (usually a CT scan) a detailed mesh is constructed. To make the information stored in this mesh available during the operation, the mesh must be registered onto the intra operative view. However, mainly due to the pneumoperitoneum, the organ has undergone important deformation between the pre-operative images acquisition and the operation. The pre-operative shape and the intra-operative shape of the organ do not correspond. Therefore a non rigid registration is required to align the mesh and the real organ. Our registration algorithms also allowed us to work on a mean to automatically recover boundary conditions of a patient specific liver.

We created a statistical atlas [29] of the human liver to store some of the liver boundary conditions positions : the veina cava and the anchor point of the falciform ligament positions. This method was presented at MICCAI 2014. We also developed a new registration method [28] that evolves automatically from a rigid registration to a non rigid registration to solve the initial alignment problem. The method use some anatomical features of the liver such as the anchor point position of the falciform ligament. This method was presented at ISBMS 2014.

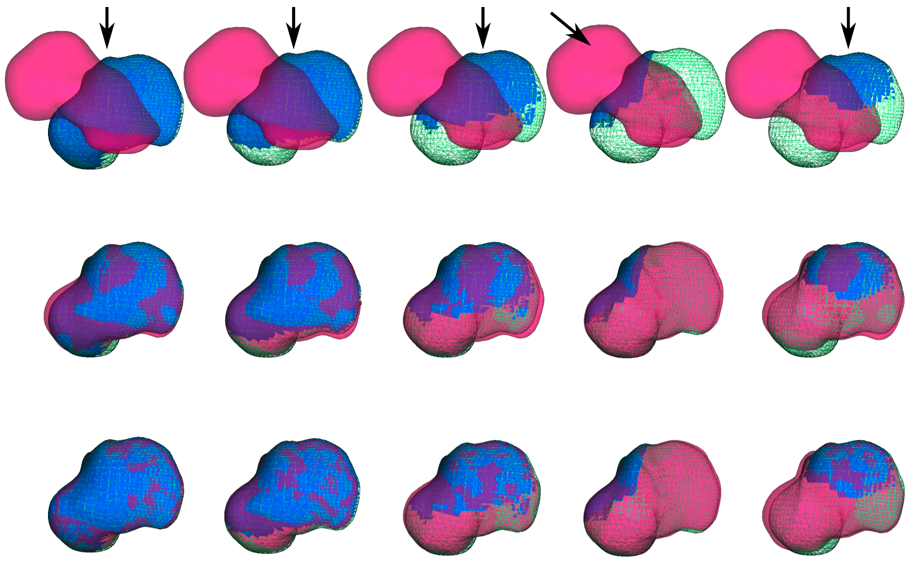


Figure 14. Results showing the initial alignment of a liver between pre-operative and intra-operative data

5.2.3.2. Augmented reality

Participant: Nazim Haouchine.

After this intra-operative registration, the augmented reality is possible. This topic is one the highlight of the year 2014. In 2014, we proposed a method for real-time augmented reality of internal liver structures during minimally invasive hepatic surgery [25]. This project is done is collaboration with the EPI MAGRIT. Vessels and tumors computed from pre-operative CT scans can be overlaid onto the laparoscopic view for surgery guidance. Compared to current methods, our method is able to locate the in-depth positions of the tumors based on partial three-dimensional liver tissue motion using a real-time biomechanical model. This model permits to properly handle the motion of internal structures even in the case of anisotropic or heterogeneous tissues, as it is the case for the liver and many anatomical structures. Experimentations conducted on phantom liver permits to measure the accuracy of the augmentation while real-time augmentation on in vivo human liver during real surgery shows the benefits of such an approach for minimally invasive surgery. Finally, a method for 3D reconstruction of elastic shapes with self-occlusion handling was also proposed.

5.2.3.3. Segmentation

Participant: Zhifan Jiang.

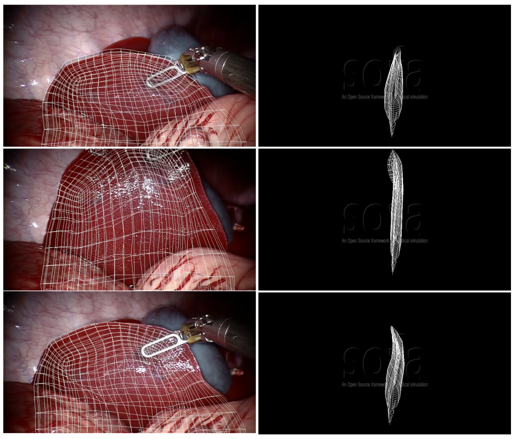


Figure 15. Augmented reality view of a liver during laparoscopic surgery

We have been working on medical image analysis in the context of the female pelvic medicine. Image-based diagnoses of pelvic floor disorders like prolapse or endometriosis rely on mechanical indicators, such as mobilities of organs and shear displacements between organs. Image data do not provide directly qualitative indicators hence analysis and diagnosis of medical are required although unfortunately subjected to surgeon expertise subjectivity. Therefore, objective information would be useful for both precise diagnoses and planning of surgical procedure. The objective is to develop numerical tools which extract quantitative information from static and cine MR images based on algorithms of detection and tracking.

We have developed numerical models not only for visualization, but also for quantitative measurements on a group of organs, such as their shapes and their relative movements. The numerical tool extracts these quantitative information (displacements and shear inter-organ) as well as the geometric shape of organs from images via Model-to-Image registration based on B-spline models. Our approach enables to identify multiple organ shapes in a single 2-dimension MR image and then to track their motion in a sequence of 2-dimension dynamic (cine) MR images for the study of the mobilities of the pelvic system. The method has been tested on healthy and pathological patient-specific data (19 patients) and the results provide valuable data to assess the shear displacement between organs and therefore making it possible to identify weakened ligaments or fascia which function differently in patients having pathologies. However, the results are to be validated by further mechanical FEM simulations. This work has been accepted in the journal STRAIN.

5.2.3.4. MIND project

Participants: Myriam Lekkal, Raffaella Trivisonne.

Within a feasibility study contest, we worked on Human Computer Interaction developing a new, intuitive and efficient way to interact with medical information in modern operating room. Nowadays operating rooms are progressively outfitted with computerized equipment necessary to access and manipulate a significant amount of data (i.e. medical images, patient's records, patient's vitals and physical parameters of the operating environment). This type of equipment belongs to the non-sterile section of an operating room, therefore surgeons, who are not allowed to be contaminated, cannot directly interact with it.

The idea of MIND project is to create a new device that could be used alone, such as a remote control, or easily integrated onto several locations, according to user preferences or constraints from the surgical procedure. Through this remote control, surgeons are able to access and manipulate medical information within the operating field and without leaving the instruments. For the software side the main aspects are distributed in two categories: a low level library, in charge of tasks such like handling the communications between

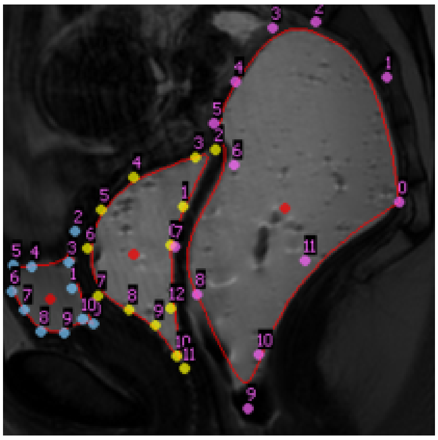
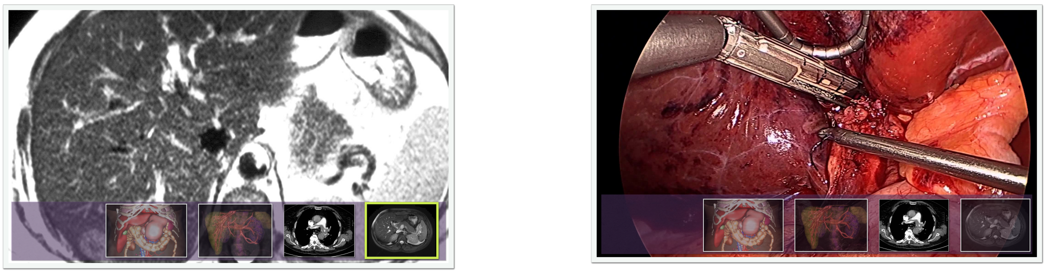


Figure 16. Contour segmentation on the pelvic system

the wireless instrument and the central computer, and a set of high-level functionalities and applications concerning the development of users GUI and new applications according to the needs of the case.

This work resulted in a patent [35] (still pending). Read more here http://mindsurgeonmouse.weebly.com/.



(a) View 1 (b) View 2

Figure 17. Example of the MIND GUI

SIERRA Project-Team

5. New Results

5.1. An Optimal Affine Invariant Smooth Minimization Algorithmn

Participant: Alexandre d'Aspremont.

We formulate an affine invariant implementation of the algorithm in Nesterov (1983). We show that the complexity bound is then proportional to an affine invariant regularity constant defined with respect to the Minkowski gauge of the feasible set. We also detail matching lower bounds when the feasible set is an ℓp ball. In this setting, our bounds on iteration complexity for the algorithm in Nesterov (1983) are thus optimal in terms of target precision, smoothness and problem dimension. (in collaboration with Cristóbal Guzmán, Martin Jaggi)

5.2. SAGA: A Fast Incremental Gradient Method With Support for Non-Strongly Convex Composite Objectives

Participants: Simon Lacoste-Julien, Francis Bach.

In this work we introduce a new optimisation method called SAGA in the spirit of SAG, SDCA, MISO and SVRG, a set of recently proposed incremental gradient algorithms with fast linear convergence rates. SAGA improves on the theory behind SAG and SVRG, with better theoretical convergence rates, and has support for composite objectives where a proximal operator is used on the regulariser. Unlike SDCA, SAGA supports non-strongly convex problems directly, and is adaptive to any inherent strong convexity of the problem. Moreover, the proof of the convergence bounds is much simpler than the one of our earlier work SAG. (in collaboration with A. Defazio, ANU)

5.3. Non-parametric Stochastic Approximation with Large Step sizes

Participants: Aymeric Dieuleveut, Francis Bach.

We consider the random-design least-squares regression problem within the reproducing kernel Hilbert space (RKHS) framework. Given a stream of independent and identically distributed input/output data, we aim to learn a regression function within an RKHS $\mathcal{H}$, even if the optimal predictor (i.e., the conditional expectation) is not in $\mathcal{H}$. In a stochastic approximation framework where the estimator is updated after each observation, we show that the averaged unregularized least-mean-square algorithm (a form of stochastic gradient), given a sufficient large step-size, attains optimal rates of convergence for a variety of regimes for the smoothnesses of the optimal prediction function and the functions in $\mathcal{H}$.

5.4. Adaptivity of averaged stochastic gradient descent to local strong convexity for logistic regression

Participant: Francis Bach.

In this work, we consider supervised learning problems such as logistic regression and study the stochastic gradient method with averaging, in the usual stochastic approximation setting where observations are used only once. We show that after N iterations, with a constant step-size proportional to $1/R^2\sqrt{N}$ where N is the number of observations and R is the maximum norm of the observations, the convergence rate is always of order $O(1/\sqrt{N})$, and improves to $O(R^2/\mu N)$ where μ is the lowest eigenvalue of the Hessian at the global optimum (when this eigenvalue is greater than $R^2/\sqrt{N}$). Since μ does not need to be known in advance, this shows that averaged stochastic gradient is adaptive to *unknown local* strong convexity of the objective function. Our proof relies on the generalized self-concordance properties of the logistic loss and thus extends to all generalized linear models with uniformly bounded features.

5.5. Serialrank: Spectral Ranking using Seriation

Participants: Fajwel Fogel, Alexandre d'Aspremont.

We describe a seriation algorithm for ranking a set of n items given pairwise comparisons between these items. Intuitively, the algorithm assigns similar rankings to items that compare similarly with all others. It does so by constructing a similarity matrix from pairwise comparisons, using seriation methods to reorder this matrix and construct a ranking. We first show that this spectral seriation algorithm recovers the true ranking when all pairwise comparisons are observed and consistent with a total order. We then show that ranking reconstruction is still exact even when some pairwise comparisons are corrupted or missing, and that seriation based spectral ranking is more robust to noise than other scoring methods. An additional benefit of the seriation formulation is that it allows us to solve semi-supervised ranking problems. Experiments on both synthetic and real datasets demonstrate that seriation based spectral ranking achieves competitive and in some cases superior performance compared to classical ranking methods. (in coolaboration with Milan Vojnovic, Microsoft Research).

5.6. Sequential Kernel Herding: Frank-Wolfe Optimization for Particle Filtering

Participants: Simon Lacoste-Julien, Francis Bach.

Recently, the Frank-Wolfe optimization algorithm was suggested as a procedure to obtain adaptive quadrature rules for integrals of functions in a reproducing kernel Hilbert space (RKHS) with a potentially faster rate of convergence than Monte Carlo integration (and "kernel herding" was shown to be a special case of this procedure). In this paper, we propose to replace the random sampling step in a particle filter by Frank-Wolfe optimization. By optimizing the position of the particles, we can obtain better accuracy than random or quasi-Monte Carlo sampling. In applications where the evaluation of the emission probabilities is expensive (such as in robot localization), the additional computational cost to generate the particles through optimization can be justified. Experiments on standard synthetic examples as well as on a robot localization task indicate indeed an improvement of accuracy over random and quasi-Monte Carlo sampling. (in collaboration with Fredrik Lindsten, Cambridge University)

5.7. Learning to Learn for Structured Sparsity

Participants: Nino Shervashidze, Francis Bach.

Structured sparsity has recently emerged in statistics, machine learning and signal processing as a promising paradigm for learning in high-dimensional settings. All existing methods for learning under the assumption of structured sparsity rely on prior knowledge on how to weight (or how to penalize) individual subsets of variables during the subset selection process, which is not available in general. Inferring group weights from data is a key open research problem in structured sparsity.

In this work, we propose a Bayesian approach to the problem of group weight learning. We model the group weights as hyperparameters of heavy-tailed priors on groups of variables and derive an approximate inference scheme to infer these hyperparameters. We empirically show that we are able to recover the model hyperparameters when the data are generated from the model, and we demonstrate the utility of learning weights in synthetic and real denoising problems.

5.8. Analysis of purely random forests bias

Participant: Sylvain Arlot.

Random forests are a very effective and commonly used statistical method, but their full theoretical analysis is still an open problem. As a first step, simplified models such as purely random forests have been introduced, in order to shed light on the good performance of random forests. In this paper, we study the approximation error (the bias) of some purely random forest models in a regression framework, focusing in particular on the influence of the number of trees in the forest. Under some regularity assumptions on the regression function, we show that the bias of an infinite forest decreases at a faster rate (with respect to the size of each tree) than a single tree. As a consequence, infinite forests attain a strictly better risk rate (with respect to the sample size) than single trees. Furthermore, our results allow to derive a minimum number of trees sufficient to reach the same rate as an infinite forest. As a by-product of our analysis, we also show a link between the bias of purely random forests and the bias of some kernel estimators. (In collaboration with Robin Genuer, Université de Bordeaux)

5.9. Large-Margin Metric Learning for Constrained Partitioning Problems

Participants: Rémi Lajugie, Sylvain Arlot, Francis Bach.

We consider unsupervised partitioning problems based explicitly or implicitly on the minimization of Euclidean distortions, such as clustering, image or video segmentation, and other change-point detection problems. We emphasize on cases with specific structure, which include many practical situations ranging from mean-based change-point detection to image segmentation problems. We aim at learning a Mahalanobis metric for these unsupervised problems, leading to feature weighting and/or selection. This is done in a supervised way by assuming the availability of several (partially) labeled datasets that share the same metric. We cast the metric learning problem as a large-margin structured prediction problem, with proper definition of regularizers and losses, leading to a convex optimization problem which can be solved efficiently. Our experiments show how learning the metric can significantly improve performance on bioinformatics, video or image segmentation problems.

5.10. Metric Learning for Aligning temporal sequences

Participants: Damien Garreau, Rémi Lajugie, Sylvain Arlot, Francis Bach.

In this work, we propose to learn a Mahalanobis distance to perform alignment of multivariate time series. The learning examples for this task are time series for which the true alignment is known. We cast the alignment problem as a structured prediction task, and propose realistic losses between alignments for which the optimization is tractable. We provide experiments on real data in the audio to audio context, where we show that the learning of a similarity measure leads to improvements in the performance of the alignment task. We also propose to use this metric learning framework to perform feature selection and, from basic audio features, build a combination of these with better performance for the alignment.

5.11. Weakly Supervised Action Labeling in Videos Under Ordering Constraints

Participants: Rémi Lajugie, Francis Bach.

We are given a set of video clips, each one annotated with an ordered list of actions, such as “walk” then “sit” then “answer phone” extracted from, for example, the associated text script. We seek to temporally localize the individual actions in each clip as well as to learn a discriminative classifier for each action. We formulate the problem as a weakly supervised temporal assignment with ordering constraints. Each video clip is divided into small time intervals and each time interval of each video clip is assigned one action label, while respecting the order in which the action labels appear in the given annotations. We show that the action label assignment can be determined together with learning a classifier for each action in a discriminative manner. We evaluate the proposed model on a new and challenging dataset of 937 video clips with a total of 787720 frames containing sequences of 16 different actions from 69 Hollywood movies. (in collaboration with Piotr Bojanowski, Ivan Laptev, Jean Ponce, Cordelia Schmid and Josef Sivic)

5.12. On Pairwise Cost for Multi-Object Network Flow Tracking

Participant: Simon Lacoste-Julien.

Multi-object tracking has been recently approached with the min-cost network flow optimization techniques. Such methods simultaneously resolve multiple object tracks in a video and enable modeling of dependencies among tracks. Min-cost network flow methods also fit well within the "tracking-by-detection" paradigm where object trajectories are obtained by connecting per-frame outputs of an object detector. Object detectors, however, often fail due to occlusions and clutter in the video. To cope with such situations, we propose an approach that regularizes the tracker by adding second order costs to the min-cost network flow framework. While solving such a problem with integer variables is NP-hard, we present a convex relaxation with an efficient rounding heuristic which empirically gives certificates of small suboptimality. Results are shown on real-world video sequences and demonstrate that the new constraints help selecting longer and more accurate tracks improving over the baseline tracking-by-detection method. (in collaboration with Visesh Chari,Ivan Laptev, Josef Sivic).

5.13. A Markovian approach to distributional semantics with application to semantic compositionality

Participants: Edouard Grave, Francis Bach, Guillaume Obozinski.

In this work, we describe a new approach to distributional semantics. This approach relies on a generative model of sentences with latent variables, which takes the syntax into account by using syntactic dependency trees. Words are then represented as posterior distributions over those latent classes, and the model allows to naturally obtain in-context and out-of-context word representations, which are comparable. We train our model on a large corpus and demonstrate the compositionality capabilities of our approach on different datasets.

5.14. A convex relaxation for weakly supervised relation extraction

Participant: Edouard Grave.

A promising approach to relation extraction, called weak or distant supervision, exploits an existing database of facts as training data, by aligning it to an unlabeled collection of text documents. Using this approach, the task of relation extraction can easily be scaled to hundreds of different relationships. However, distant supervision leads to a challenging multiple instance, multiple label learning problem. Most of the proposed solutions to this problem are based on non-convex formulations, and are thus prone to local minima. In this article, we propose a new approach to the problem of weakly supervised relation extraction, based on discriminative clustering and leading to a convex formulation. We demonstrate that our approach outperforms state-of-the-art methods on a challenging dataset introduced in 2010.

5.15. Weakly supervised named entity classification

Participant: Edouard Grave.

In this paper, we describe a new method for the problem of named entity classification for specialized or technical domains, using distant supervision. Our approach relies on a simple observation: in some specialized domains, named entities are almost unambiguous. Thus, given a seed list of names of entities, it is cheap and easy to obtain positive examples from unlabeled texts using a simple string match. Those positive examples can then be used to train a named entity classifier, by using the PU learning paradigm, which is learning from positive and unlabeled examples. We introduce a new convex formulation to solve this problem, and apply our technique in order to extract named entities from financial reports corresponding to healthcare companies.

5.16. Fast imbalanced binary classification: a moment-based approach

Participant: Edouard Grave.

In this paper, we consider the problem of imbalanced binary classification in which the number of negative examples is much larger than the number of positive examples. The two mainstream methods to deal with such problems are to assign different weights to negative and positive points or to subsample points from the negative class. In this paper, we propose a different approach: we represent the negative class by the two first moments of its probability distribution (the mean and the covariance), while still modeling the positive class by individual examples. Therefore, our formulation does not depend on the number of negative examples, making it suitable to highly imbalanced problems and scalable to large datasets. We demonstrate empirically, on a protein classification task and a text classification task, that our approach achieves similar statistical performance than the two mainstream approaches to imbalanced classification problems, while being more computationally efficient. (in collaboration with Laurent El Ghaoui, U.C. Berkeley)

SIROCCO Project-Team

6. New Results

6.1. Analysis and modeling for compact representation and navigation

3D modelling, multi-view plus depth videos, Layered depth images (LDI), 2D and 3D meshes, epitomes, image-based rendering, inpainting, view synthesis

6.1.1. Salient object detection

Participants: Olivier Le Meur, Zhi Liu.

Salient object detection consists in extracting in an automatic manner the most interesting object in an image or video sequence. From an input image, an object, with well-defined boundaries, is detected based on its saliency. This subject knows an renewed interest these last years. A number of datasets serving as ground truth has been released and can be used to benchmark methods.

In 2013, we proposed a new method for detecting salient objects in still color images. In 2014, this method has been extended to video sequences [21]. Based on the superpixel representation of video frames, motion histograms and color histograms are computed at local and global levels. From these histograms, a superpixel-level temporal saliency measure as well as a spatial saliency measure are obtained. Finally, a pixel-level saliency derivation method is proposed to generate pixel-level temporal saliency map and spatial saliency map. An adaptive fusion method allows to integrate them into an unique spatiotemporal saliency map. Experimental results on two public datasets demonstrate that the proposed model outperforms state-of-the-art spatiotemporal saliency model in terms of both saliency detection and human fixation prediction.

6.1.2. Saliency aggregation

Participants: Olivier Le Meur, Zhi Liu.

In this study [32], we investigate whether the aggregation of saliency maps allows to outperform the best saliency models. Today there exist a number of saliency models for predicting the most visually salient locations within a scene. Although all existing models follow the same objective, they provide results which could be, to some extent, different. The discrepancies are related to the quality of the prediction but also to the saliency map representation. Indeed some models output very focused saliency maps whereas the distribution of saliency values is much more uniform in other models. Others tend to emphasize more on the image edges, the color or luminance contrast. This saliency map manifold contains a rich resource that should be used and from which new saliency maps could be inferred. Combining saliency maps generated using different models might enhance the prediction quality and the robustness of the prediction. Our goal is then to take saliency maps from this manifold and to produce the final saliency map.

This study discussed various aggregation methods; six unsupervised and four supervised learning methods are tested on two existing eye fixation datasets. Results show that a simple average of the TOP 2 saliency maps significantly outperforms the best saliency models. Considering more saliency models tends to decrease the performance, even when robust aggregation methods are used. Concerning the supervised learning methods, we provide evidence that it is possible to further increase the performance, under the condition that an image similar to the input image can be found in the training dataset. Our results might have an impact for critical applications which require robust and relevant saliency maps.

6.1.3. Models for 3D video quality assessment

Participants: Darya Khaustova, Olivier Le Meur.

This work is carried out in collaboration with Orange labs. The goal is to design objective metrics for quality assessment of 3D video content, by establishing links between human visual perception (visual comfort) and video parameters such as quality and depth quantity, and between visual comfort and visual attention. In 2013 we investigated the differences in 2D visual attention in comparison with 3D visual attention [31]. In 2014, we have focused on the design of an objective stereoscopic quality metric. In stereoscopic video quality, the assessment of spatial and temporal distortions by conventional quality metrics became incomplete because of the added depth dimension. Improperly captured or rendered, depth information can induce visual discomfort, impacting the overall video 3D QoE quality independently of image quality. The model is based on perceptual thresholds, namely visual annoyance, and acceptability. The visual annoyance threshold defines the boundary between annoying and not annoying sensation: 50% of subjects consider a stimulus annoying and 50% as not annoying. Acceptability determines the viewer's expectation level for the perceived video quality in a certain context and situation (inspired by the acceptability for the customer defined as an adequate service.

In order to compute the quality score, the proposed metric requires in input the distortion level of a technical and particular parameter, annoyance threshold and acceptability threshold of the targeted parameter. The performance of proposed objective mode is evaluated by considering five view asymmetries with five degradation levels. Generated contents were assessed by 30 subjects for each asymmetry (focal length mismatch, vertical shift, and rotation, green and white level reduction). The results of the subjective test have demonstrated that it is possible to classify detected problem to one of the objective categories using corresponding acceptability and visual annoyance thresholds.

6.1.4. Epitome-based video representation

Participants: Martin Alain, Christine Guillemot.

In 2014, we have developed fast methods for constructing epitomes from images. An epitome is a factorized texture representation of the input image, and its construction exploits self-similarities within the image. Known construction methods are memory and time consuming. The proposed methods, using dedicated list construction on one hand and clustering techniques on the other hand, aim at reducing the complexity of the search for self-similarities. Experiments show that interesting complexity results can be obtained without degrading the epitome quality for both proposed methods. By limiting the number of exhaustive searches we limit the memory occupation and the processing time, while keeping a good epitome quality (down to 18.08 % of the original memory occupation and 41.39 % of the original processing time) [25]. As an example, images reconstructed using the different techniques are visible in Fig. 1 . The epitome construction method is currently being extended from still images to groups of images in video sequences. Denoising and super-resolution algorithms based on the constructed epitomes are also under study.

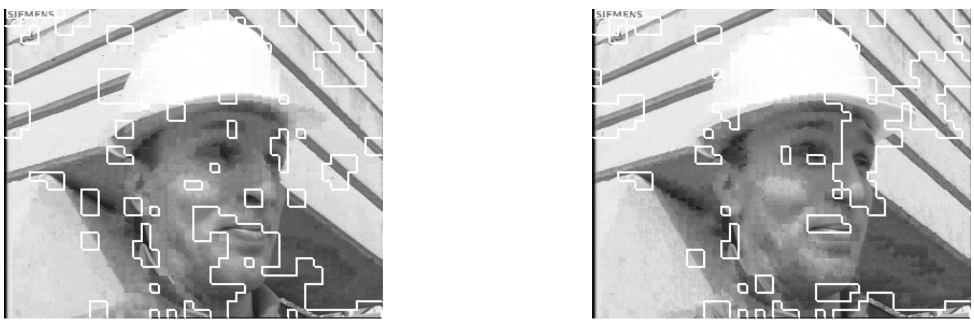


Figure 1. Reconstructed images using the list-based (left) and clustering-based methods. Epitome patches are highlighted in white.

6.1.5. Light field tomographic reconstruction from a fixed camera focal stack

Participants: Christine Guillemot, Elif Vural.

Thanks to the internship of Antoine Mousnier (student at Ecole Centrale Lyon), we have developed a novel approach to partially reconstruct high-resolution 4D light fields from a stack of differently focused photographs taken with a fixed camera. First, a focus map is calculated from this stack using a simple approach combining gradient detection and region expansion with graph cut. Then, this focus map is converted into a depth map thanks to the calibration of the camera. We proceed after this with the tomographic reconstruction of the epipolar images by back-projecting the focused regions of the scene only. We call it masked back-projection. The angles of back-projection are calculated from the depth map. Thanks to the high angular resolution we achieve, we are able to render puzzling perspective shifts although the original photographs were taken from a single fixed camera at a fixed position and render images with extended focus (see Fig. 2). To the best of our knowledge, our method is the first one to reconstruct a light field by using a focal stack captured with an ordinary camera at a fixed viewpoint.

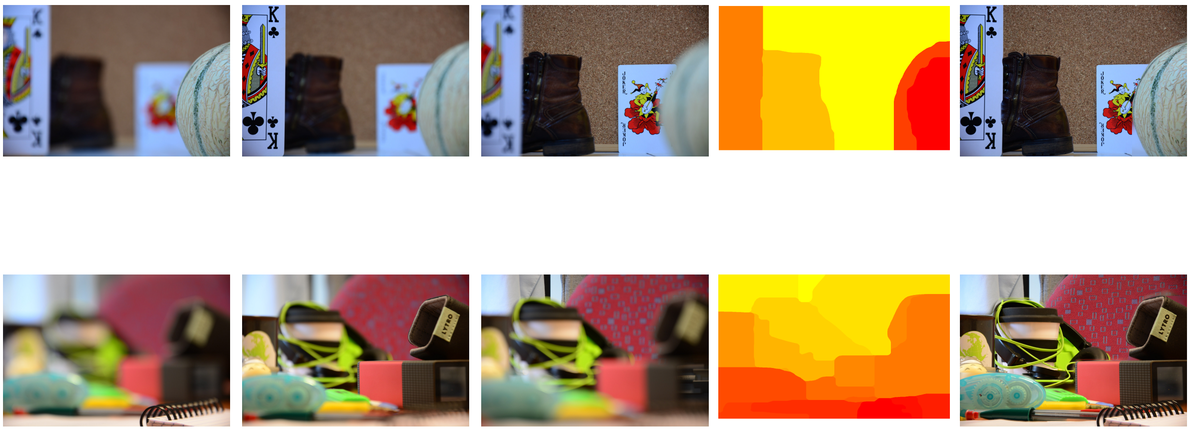


Figure 2. Three images of the focal stack (left); estimated depth map and image with extended focus (right). The focal stack images of the first and second rows have been captured with a Nikon 5200 camera.

6.2. Rendering, inpainting and super-resolution

image-based rendering, inpainting, view synthesis, super-resolution

6.2.1. Video inpainting

Participants: Mounira Ebdelli, Christine Guillemot, Olivier Le Meur.

Image (and video) inpainting refers to the process of restoring missing or damaged areas in an image (or a video). This field of research has been very active over the past years, boosted by numerous applications: restoring images from scratches or text overlays, loss concealment in a context of impaired image transmission, object removal in a context of editing, disocclusion in image-based rendering of viewpoints different from those captured by the cameras. Inpainting is an ill-posed inverse problem: given observations, or known samples in a spatial (or spatio-tempoal) neighborhood, the goal is to estimate unknown samples of the region to be filled in. Many methods already exist for image inpainting, either based on PDE (Partial Derivative Equation)-based diffusion schemes, either using sparse or low rank priors or following texture synthesis principles exploiting statistical or self-similarity priors.

In 2014, the problem of video inpainting has been further addressed with free-moving cameras. The algorithm developed first compensates the camera motion between the current frame and its neighboring frames in a sliding window, using a new region-based homography computation which better respects the geometry of the scene compared to state-of-the-art methods. The source frame is first segmented into regions in order to find homogeneous regions. Then, the homography for mapping each region into the target frame is estimated. The overlapping of all aligned regions forms the registration of the source frame into the target one. Once the neighboring frames have been aligned, they form a stack of images from which the best candidate pixels are searched in order to replace the missing ones. The best candidate pixel is found by minimizing a cost function which combines two energy terms. One energy term, called the data term, captures how stationary is the background information after registration, hence enforcing temporal coherency. The second term aims at favoring spatial consistency and preventing incoherent seams, by computing the energy of the difference between each candidate pixel and its 4-neighboring pixels in the missing region. The minimization of the energy term is performed globally using Markov Random Fields and graph cuts. A method of Poisson blending has been implemented in order to further enhance the visual quality of the inpainted videos. The proposed approach, although less complex than state-of-the-art methods, provides more natural results.

6.2.2. Image and video super-resolution in the example-based framework

Participants: Marco Bevilacqua, Christine Guillemot, Aline Roumy.

Super-resolution (SR) refers to the problem of creating a high-resolution (HR) image, given one or multiple low-resolution (LR) images as input. The SR process aims at adding to the LR input(s) new plausible high frequency details, to a greater extent than traditional interpolation methods. We mostly focused on the single-image problem, where only a single LR image is available. We have adopted the example-based framework on one hand and the sparse approximation framework on the other hand.

In the example-based framework, the relation between the LR and HR image spaces is modeled with the help of pairs of small "examples", i.e. texture patches. Each example pair consists of a LR patch and its HR version that also includes high-frequency details; the pairs of patches form a dictionary of patches. For each patch of the LR input image, one or several similar patches are found in the dictionary, by performing a nearest neighbor search. The corresponding HR patches in the dictionary are then combined to form a HR output patch; and finally all the reconstructed HR patches are re-assembled to build the superresolved image. In this procedure, one important aspect is how the dictionary of patches is built. At this regard, two choices are possible: an external dictionary, formed by sampling HR and LR patches from external training images; and an internal dictionary, where the LR/HR patch correspondences are learned by putting in relation directly the input image and scaled versions of it. The advantage of having an external dictionary is that it is built in advance: this leads to a reduction of the computational time, whereas in the internal case the dictionary is generated online at each run of the algorithm. However, external dictionaries have a considerable drawback: they are fixed and so non-adapted to the input image. To be able to satisfactorily process any input image, we need then to include in the dictionary a large variety of patch correspondences, leading to a high computational time. In 2013, external dictionaries have been designed to bridge the gap between external and internal dictionary based methods.

In 2014 instead, we proposed a novel SR method for internal dictionaries [16]. The internal dictionary contains pair of LR/HR patches taken from the image to be processed and is by construction well adapted to the data. However, its size is limited since it results from the sampling of a single image. This leads to an undersampling of the LR space and even more of the HR space. To overcome this problem, state of the art methods select, for each input LR patch, a local neighborhood, learn the local geometry of this neighborhood, and apply it in the HR domain. Therefore, an underlying hypothesis is that the local neighborhoods in the LR and HR domain are similar. To avoid this hypothesis, we employ a regression-based method to directly map LR input patches into their related HR output patches. To make this regression more robust, first the LR patches have been first oversampled (by a bicubic interpolation) such that LR and HR spaces have the same dimension, and second a Tikhonov regularization has been added. When compared to other state-of-the-art algorithms, our proposed algorithm shows the best performance, both in terms of objective metrics and subjective visual results. As for the former, it presents considerable gains in PSNR and SSIM values. When observing the super-resolved images, also, it turns out to be the most capable in producing fine artifact-free HR details.

6.2.3. Image super-resolution in a sparse and manifold learning framework

Participants: Julio Cesar Ferreira, Christine Guillemot, Olivier Le Meur, Elif Vural.

The problem of image super-resolution has also been addressed in a sparse approximation framework. This led to a novel algorithm based on sparse representations in which a structure tensor-based regularization has been introduced [29]. The relative discrepancy between the two eigenvalues of the structure tensor is an indicator of the degree of anisotropy of the gradient in a region of the image. The eigenvalues and eigenvectors of the structure tensor are used to compute, for each pixel belonging to a salient edge, a stream line in the direction perpendicular to the edge (given by the eigenvector corresponding to the highest eigenvalue of the structure tensor). The saliency of an edge is given by the S-norm of the highest eigenvalue. An energy term dealing with the sharpness of edges is then computed and used as a regularization constraint to modify the current estimated high resolution image inside the Iterative Shrinkage Thresholding algorithm. This extra constraint forces the value of the current pixel along the stream line to be as close as possible to pixel values having lowest saliency. The resulting single-image algorithm, called Sharper Edges based Adaptive Sparse Domain Selection (SE-ASDS) allows sharpening edges and reducing the ringing artefacts compared to existing methods. This is illustrated in Fig.3

Figure 3. Comparison of SR results ($\times 3$). (a) LR image;(b) Nearest-neighbor; (c) Sparse method without structure-based regularization; (d) SE-ASDS results. (e) Comparison between (c) and (d) on patches: edges of (d) are more contrasted than (c).

In the previous method, the dictionaries used for the sparse approximation method are defined as a union of PCA basis learned on clusters of patches of the input image. The clusters are constructed using the classical k-means algorithm with patch distances computed with the Euclidean distance. This study is being pursued by assuming manifold models for the patches of the input images. A method using graph-based clustering has then been used for clustering patches on the manifold, and this method has been extended to cope with the out-of-sample problem. Dedicated dictionary learning methods are currently under development to have dictionaries best adapted to the manifold structure.

6.3. Representation and compression of large volumes of visual data

Sparse representations, data dimensionality reduction, compression, scalability, perceptual coding, rate-distortion theory

6.3.1. Manifold learning and low dimensional embedding for classification

Participants: Christine Guillemot, Elif Vural.

Typical supervised classifiers such as SVM are designed for generic data types and do not make any particular assumption about the geometric structure of data, while data samples have an intrinsically low-dimensional structure in many data analysis applications. Recently, many supervised manifold learning methods have been proposed in order to take the low-dimensional structure of data into account when learning a classifier. Unlike unsupervised manifold learning methods which only take the geometric structure of data samples into account when learning a low-dimensional representation, supervised manifold learning methods learn an embedding that not only preserves the manifold structure in each class, but also enhances the separation between different classes.

An important factor that influences the performance of classification is the separability of different classes in the computed embedding. We thus do a theoretical analysis of separability of data representations given by supervised manifold learning. In particular, we focus on the nonlinear supervised extensions of the Laplacian eigenmaps algorithm and examine the linear separation between different classes in the learned embedding. We first consider a setting with two classes and show that the two classes become linearly separable even with a one-dimensional embedding. We characterize the linear separation in terms of the data graph properties such as edge weights, diameter, and volume and some algorithm parameters. We then extend these results to a setting with multiple classes, where the classes are assumed to be categorizable into a few groups with high intra-group affinities. We show that, if the graph is such that the inter-group graph weights are sufficiently small, the learned embedding becomes linearly separable at a dimension that is proportional to the number of groups. These theoretical findings are also confirmed by experimentation on synthetic data sets and image data.

Next, we consider the problem of out-of-sample generalizations for manifold learning. Most manifold learning methods compute an embedding in a pointwise manner, i.e., data coordinates in the learned domain are computed only for the initially available training data. The generalization of the embedding to novel data samples is an important problem, especially in classification problems. Previous works for out-of-sample generalizations are designed for unsupervised methods. We study the problem for the particular application of data classification and propose an algorithm to compute a continuous function from the original data space to the low-dimensional space of embedding. In particular, we construct an interpolation function in the form of a radial basis function that maps input points as close as possible to their projections onto the manifolds of their own class. Experimental results show that the proposed method gives promising results in the classification of low-dimensional image data such as face images.

6.3.2. Dictionary learning for sparse coding and classification of satellite images

Participants: Jeremy Aghaei Mazaheri, Christine Guillemot, Claude Labit.

In the context of the national partnership Inria-Astrium, we explore novel methods to encode images captured by a geostationary satellite. These pictures have to be compressed on-board before being sent to earth. Each picture has a high resolution and so the rate without compression is very high (about 70 Gbits/sec) and the goal is to achieve a rate after compression of 600 Mbits/sec, that is a compression ratio higher than 100. On earth, the pictures are decompressed with a high reconstruction quality and visualized by photo-interpreters. The goal of the study is to design novel transforms based on sparse representations and learned dictionnaries for satellite images.

We have developed methods for learning adaptive tree-structured dictionaries. Each dictionary in the structure is learned on a subset of residuals from the previous level, with the K-SVD algorithm. The tree structure offers better rate-distortion performance than a "flat" dictionary learned with K-SVD, especially when only a few atoms are selected among the first levels of the tree. The tree-structured dictionary allows efficient coding of the indices of the selected atoms. Besides coding, these structured dictionaries turn out to be useful tools for

MTF (Modulation Transfer Function) estimation and supervised classification. The MTF estimation consists in estimating the MTF of the instrument used to take this picture. The learned structured dictionaries are currently studied to perform supervised classification in a context of scene recognition in satellite images. In that case, dictionaries should be learned for specific scenes. Then, patches (around each pixel) of a test picture to classify are decomposed over the different dictionaries to determine for each pixel the dictionary giving the best approximation and thus the corresponding class. A graph-cut algorithm can be applied to smooth the classification results. We are currently trying to learn more discriminant dictionaries for this specific application. For that purpose, the objective function to minimize to learn the dictionaries should not only be reconstructive, but also discriminative.

6.3.3. Adaptive clustering with Kohonen self-organizing maps for second-order prediction

Participants: Christine Guillemot, Bihong Huang.

The High Efficiency Video Coding standard (HEVC) supports a total of 35 intra prediction modes which aim at reducing spatial redundancy by exploiting pixel correlation within a local neighborhood. However the correlation remains in the residual signals of intra prediction, leading to some high energy prediction residuals. In 2014, we have studied several methods to exploit remaining correlation in residual domain after intra prediction. The method uses vector quantization with codebooks learned and dedicated to the different prediction modes in order to model the directional characteristics of the residual signals. The best matching code vector is found in a rate-distortion optimization sense. Finally, the index of the best matching code vector is sent to the decoder and the vector quantization error, the difference between the intra residual vector and the best matching code vector, is processed by the conventional operations of transform, scalar quantization and entropy coding. In a first approach, the codebooks are learned using the k-means algorithm. The learning algorithm proceeds in two passes so that the training set of residual vectors corresponds to the case where the vector quantization is the best mode in rate-distortion sense for the second-order prediction. It has been observed that the codebooks learned for different Quantization Parameters (QP) are very similar, leading eventually to QP-independent codebooks. A second method is being developed using clustering with Kohonen self-organizing maps in the codebook learning stage.

6.3.4. HDR video compression

Participants: Christine Guillemot, Mikael Le Pendu.

High Dynamic Range (HDR) images contain more intensity levels than traditional image formats. Instead of 8 or 10 bit integers, floating point values requiring much higher precision are used to represent the pixel data, leading to new compression challenges. In collaboration with Technicolor, we have developed a method for converting the floating point RGB values to high bit depth integers with an approximate logarithmic encoding that is reversible without loss. This bit depth reduction is performed adaptively depending on the minimum and maximum values which characterize the dynamic of the data. A 50% rate saving has been obtained at high bitrates compared to the well-known adaptive LogLuv transform [33]. A reversible tone mapping-operator (TMO) has also been designed for efficient compression of High Dynamic Range (HDR) images using a Low Dynamic Range (LDR) encoder. Based on a statistical model of the HDR compression scheme and assumptions on the rate of the encoded LDR image, a closed form solution has been derived for the optimal tone curve in a rate-distortion sense [34].

6.3.5. HEVC-based UHD video coding optimization

Participants: Nicolas Dhollande, Christine Guillemot, Olivier Le Meur.

The HEVC (High Efficiency Video Coding) standard brings the necessary quality versus rate performance for efficient transmission of Ultra High Definition formats (UHD). However, one of the remaining barriers to its adoption for UHD content is the high encoding complexity. We address the problem of HEVC encoding complexity reduction by proposing a strategy to infer UHD coding modes and quadtree from those optimized on the lower (HD) resolution version of the input video. A speed-up by a factor of 3 is achieved compared to directly encoding the UHD format at the expense of a limited PSNR-rate loss [28]. Another method which is still under investigation is to extract from the input video sequence a number of low-level features for adapting the coding decision such as the decomposition of the quadtree. The low-level features are related to gradient-based statistics, structure tensors statistics or entropy etc.

6.4. Distributed processing and robust communication

Information theory, stochastic modelling, robust detection, maximum likelihood estimation, generalized likelihood ratio test, error and erasure resilient coding and decoding, multiple description coding, Slepian-Wolf coding, Wyner-Ziv coding, information theory, MAC channels

6.4.1. Universal distributed source coding

Participant: Aline Roumy.

In 2012, we started a new collaboration with Michel Kieffer and Elsa Dupraz (Supelec, L2S) on universal distributed source coding. Distributed source coding (DSC) refers to the problem where several correlated sources need to be compressed without any cooperation at the encoders. Decoding is however performed jointly. This problem arises in sensor networks but also in video compression techniques, where the successive frames are seen as distributed such that the correlation between the frames is not directly used at the encoder. Traditional approaches for DSC (from an information theoretical but also practical point of view) assume that the joint distribution of the sources is perfectly known. Since this assumption is not satisfied in practice, a way to get around this is to use a feedback channel (from the decoder to the encoder), that can trigger the encoder. Instead, we consider universal distributed source coding, where the joint source distribution is unknown.

More precisely, we considered the problem of compressing one source, while a second source, called side information, is available at the decoder. Further, we assumed that the conditional distribution of the side information given the source is unknown at both encoder and decoder. First, we proposed in [18] four uncertainty models for this conditional distribution, and derived the information theoretical bounds. These models differ through the (partial) knowledge on the distribution the user has. This partial knowledge includes the variation speed (slow/fast), the set of possible distributions, and eventually, some a priori distribution on the class of distributions. A complete coding scheme has also been proposed that works well for any distribution in the class. At the encoder, the proposed scheme encompasses the determination of the coding rate and the design of the encoding process. These determinations directly result from the information-theoretical compression bounds. Then a novel decoder is proposed that jointly estimate the source symbols and the conditional distribution. As the proposed decoder is based on the Expectation-Maximization algorithm, which is very sensitive to initialization, we also propose a method to produce first a coarse estimate of the distribution. The proposed scheme avoids the use of a feedback channel or the transmission of a learning sequence, which both result in a rate increase at finite length. Moreover, the proposed algorithm use non-binary LDPC codes, such that the usual binarization of the source, which induce compression inefficiency, can be avoided.

6.4.2. Rate Distortion analysis of Compressed sensing and distributed Compressed sensing

Participant: Aline Roumy.

In collaboration with Enrico Magli and Giulio Coluccia (Polito, Torino, Italy), we studied Compressed sensing as a communication tool. Compressed sensing (CS) is an efficient acquisition scheme, where the data are projected onto a randomly chosen subspace to achieve data dimensionality reduction. The projected data are called measurements. The reconstruction is performed from these measurements, by solving underdetermined linear systems under a sparsity a priori constraint. However, the obtained measurements are reals, and therefore require an infinite precision representation. Therefore, using CS as a compression tool (in the information theoretical sense), requires to determine the trade-off between the rate necessary to encode the measurements and the distortion obtained on the data. In [17], we derive the rate-distortion (RD) function of CS and distributed CS, under the assumption that the sparsity support is perfectly known at the decoder. This provides a lower bound for any practical reconstruction algorithm.

The proof technique developed in [17] has application beyond information theory. It also provides novel analyses of CS reconstruction algorithms [27]. Classical performance analysis of reconstruction algorithms, rely on parameters that are difficult to compute (RIP, coherence of the measurement matrix), for which bounds are used. Instead, we derive exact characterization, by performing either averaged (over the measurement matrix) or asymptotic (in the size of the data) analysis.

SISTM Team

6. New Results

6.1. Highlights of the Year

A work (described below), in collaboration with M. Davis and R. Tibshirani from Standford University, has been published in the "Proceedings of the National Academy of Sciences" : [8].

Females have generally more robust immune responses than males for reasons that are not well-understood. Here we used a systems analysis to investigate these differences by analyzing the neutralizing antibody response to a trivalent inactivated seasonal influenza vaccine (TIV) and a large number of immune system components, including serum cytokines and chemokines, blood cell subset frequencies, genome-wide gene expression, and cellular responses to diverse in vitro stimuli, in 53 females and 34 males of different ages. We found elevated antibody responses to TIV and expression of inflammatory cytokines in the serum of females compared with males regardless of age. This inflammatory profile correlated with the levels of phosphorylated STAT3 proteins in monocytes but not with the serological response to the vaccine. In contrast, using a machine learning approach, we identified a cluster of genes involved in lipid biosynthesis and previously shown to be up-regulated by testosterone that correlated with poor virus-neutralizing activity in men. Moreover, men with elevated serum testosterone levels and associated gene signatures exhibited the lowest antibody responses to TIV. These results demonstrate a strong association between androgens and genes involved in lipid metabolism, suggesting that these could be important drivers of the differences in immune responses between males and females.

6.2. Analysis of purely random forests bias

In collaboration with S. Arlot, we write a research report on some theoretical results about random forests : [30].

Random forests are a very effective and commonly used statistical method, but their full theoretical analysis is still an open problem. As a first step, simplified models such as purely random forests have been introduced, in order to shed light on the good performance of random forests. In this paper, we study the approximation error (the bias) of some purely random forest models in a regression framework, focusing in particular on the influence of the number of trees in the forest. Under some regularity assumptions on the regression function, we show that the bias of an infinite forest decreases at a faster rate (with respect to the size of each tree) than a single tree. As a consequence, infinite forests attain a strictly better risk rate (with respect to the sample size) than single trees. Furthermore, our results allow to derive a minimum number of trees sufficient to reach the same rate as an infinite forest. As a by-product of our analysis, we also show a link between the bias of purely random forests and the bias of some kernel estimators.

SISYPHE Project-Team

5. New Results

5.1. Fault detection and localization in networks of transmission lines

Participants: Mohamed Oumri, Michel Sorine.

Some results have been obtained in collaboration with Florent Loete (LGEP) and Qinghua Zhang (Inria, I4S):
- *Experimental validation of the inverse scattering method for distributed characteristic impedance estimation.* Our theoretic results and numerical simulations have shown the ability of inverse scattering-based methods to diagnose soft faults in electric cables, in particular, faults implying smooth spatial variations of cable characteristic parameters. We have obtained laboratory experiments confirming the ability of the inverse scattering method for retrieving spatially distributed characteristic impedance from reflectometry measurements. Various smooth or stepped spatial variations of characteristic impedance profiles have been tested. The tested electric cables are CAN unshielded twisted pairs used in trucks and coaxial cables [37].

- *Diagnosis of networks using tagged electric lines.* A new electromagnetic marking method of transmission lines has been proposed for diagnosis of electric networks when conditions of uniqueness of the solution are not fulfilled (e.g. in case of symmetries): small non-interfering characteristic defaults are added to the lines and used as tags. A patent application has been submitted [36].

A new application of our monitoring technique has been explored in collaboration with EDF and a first result has been obtained:
- *Monitoring of post-tensioned ducts or water content in concrete walls with embedded transmission lines.* We have presented an electromagnetic method of diagnosis based on frequency domain reflectometry (FDR) associated with our inversion algorithm, ISTL (Inverse Scattering for Transmission Lines). ISTL allows one to estimate the spatial profile of the electrical impedance of the line from the FDR measurements. Experimental results on two mockups of external post-tensioned ducts with filling defects show the feasibility of the method. We will try to show the similarities between auscultation external post-tensioned ducts and measurement of water content by TDR probes (Time Domain Reflectometry) [34].

- *Fault diagnosis of wired electric networks by reflectometry.* A first extension to Baum-Liu-Tesche equations has been proposed in [31].

5.2. Cardiovascular signal processing and applications

Participants: Lisa Guigue, Claire Médigue, Michel Sorine, Serge Steer.

See the Sofware section 4.1 for a description of tools developed for *Cardiovascular Waves Analysis.*

5.3. Glycemic control in ICUs

Participant: Michel Sorine.

The results of statistical analysis of the data gathered during the large clinical trial CGAO-REA have been published in [14]: “Tight computerized versus conventional glucose control in the ICU: a randomized controlled trial”. Despite the increase in the incidence of severe hypoglycemia in our experimental group, based on the absence of difference in mortality between patients on tight computerized glucose control and those on less stringent glucose control without computerized decision support systems (CDSS), this study could pave the way for future randomized controlled trials assessing new generation CDSSs allowing the safe implementation of blood glucose control in the ICU that take into account the complexity of glucose control throughout the ICU stay and the variability of individualized insulin needs. Some new objectives for computer aided glycemic control in ICUs have been proposed in [32]. An article proposing a more detailed statistical analysis of the severe hypoglycemic events has been submitted.

5.4. Modeling and optimizing patient pathways in hospital

Participants: James Leifer, Michel Sorine.

External scientific collaboration with:
- Niccolo Curatolo, Directeur des opérations, Hôpitaux universitaires Paris-Sud, Assistance publique-Hôpitaux de Paris (AP-HP);
- Dr Maurice Raphaël, Chef de service, Urgences adultes, Hôpital Bicêtre, AP-HP;
- Dr Christophe Vincent-Cassy, Responsable des systèmes informatiques des urgences, AP-HP;
- Lucie Gaillardot-Roussel, Ingénieur en organisation, AP-HP;
- Dr Paul Jarvis, Senior consultant doctor in emergency medicine, Calderdale and Huddersfield National Health Service Foundation Trust, UK.

In 2014, we began a case study of the emergency department (ED) at Bicêtre Hospital, a large ED handling 50,000 patient visits per year, which is amongst the top 10 by volume and by annual volume growth for EDs in the Paris region.

Rather than presume the appropriateness of a predetermined scientific formalism, our strategy was to allow the application to frame a series of questions in order to lead us to experiment with several potential scientific tools at the present “low risk, high uncertainty” phase of investigation:

- *Top-down modeling:* Can we capture the expert knowledge of doctors and nurses as to the pathways followed by their patients by transforming this knowledge into a series of “use case” rules borrowed from the techniques of software specification? Can these rules by transformed into an executable model using business process modeling languages and tools (Orc, YAWL, ...) for simulating the complex parallel composition of man-machine processes in a hospital setting?

- *Bottom-up modeling:* How can the hospital be instrumented for cheaply and accurately capturing its real activity (movement of people and machines, delays, errors, ...) and tuning the parameters of the model? Can we intercept HL7 messages (a standardized electronic message format for medical data) and/or access raw time-stamped database entries to use machine learning techniques (particularly process mining) to extract from the running hospital the graphs representing the actual sequence of care events in order to get rapid feedback about the most heavily used and most often delayed path segments?

- *Underlying cost semantics:* Can we formalize in process calculi (for example, a variation of pi calculus) the “micro internal economy” of costs exchanged inside the hospital to quantify the economic performance of each patient pathway?

- *Offline experimentation and optimization:* Can potential optimization to the model be explored offline in a sort of “serious game” to allow non-intrusive experimentation with different strategies for eliminating bottlenecks, increasing flow rates, decreasing costs, etc.?

- *Data visualization for medical personnel:* Given that the medical personnel themselves are best suited to fixing the daily frictional time losses that most are resigned to accept as “part of the job”, how can the model be presented in a visually lucid manner to render the previously “invisible” aspects of the hospital's organization visible?

- *Online real-time control:* Can the feedback loop be completed and the model be used to directly provide real-time visual feedback to the hospital personnel to enable them to measure their systemic progress (or systemic unintended consequences) of their localized optimizations?

SMIS Project-Team

6. New Results

6.1. Flash-Based Data Management

Participants: Nicolas Anciaux, Matias Bjørling, Philippe Bonnet, Luc Bouganim [correspondent], Niv Dayan, Saliha Lallali, Philippe Pucheral, Iulian Sandu Popa.

There is a long tradition of work around the understanding and optimization of NAND Flash memory in the team (e.g., [7], [9]). Current work in this area covers the optimization of SSD use in DBMS engines and the design of Flash-based indexing techniques for textual and spatio-temporal data. These works on Flash-Based indexing complete the work initiated in the last years on the storage and indexing engine of PlugDB (not repeated in this report but the interested reader is referred to a DAPD'14 journal publication detailing these techniques [14]).

Flash storage optimization. Solid State Drives (SSDs), based on flash chips, are now the secondary storage of choice for data intensive applications. Database systems can now rely on high performance SSDs to store log, indexes and data either on servers or in the cloud. While SSDs provide increasingly high performance out of the box, maintaining high throughput and low latency as the utilization of SSDs increases and despite abrupt changes in the workload remains a challenge. This question is central for database designers and administrators, cloud service providers, and SSD constructors. The answer depends on write-amplification, i.e., garbage collection overhead. More specifically, the answer depends on how write-amplification evolves in time. We derived a mathematical expression that relates over provisioning to write-amplification. We introduced a new block manager, called Wolf, or WOrkload Leveler for Flash. Wolf is able to detect and quickly adapt to changes in workload by pro-actively reallocating over-provisioned space among the groups based on their changing needs. It adapts better to stable workloads by measuring the update frequencies of groups instead of making assumptions about them. It uses a novel near-optimal closed-form expression to allocate over-provisioned space to groups.

Flash-based keyword indexing. As smart objects gain the capacity to acquire, store and process large volumes of data, new services emerge. However, the smart objects have to be endowed with typical data management capabilities to enable all these services. In this work, we revisit the traditional problem of information retrieval queries over large collections of files in an embedded context. A file can be any form of document, picture or data stream associated with a set of terms. A query can be any form of keyword search using a ranking function (e.g., TF-IDF) identifying the top-k most relevant files. The proposed search engine can be used in sensors to search for relevant objects in their surroundings, in cameras to search pictures by using tags, in personal smart dongles to secure the querying of documents and files hosted in an untrusted Cloud or in smart meters to perform analytic tasks (i.e., top-k queries) over sets of events (i.e., terms) captured during time windows (i.e., files) [21]. Designing such embedded search engine is however challenging due to a combination of severe and conflicting hardware constraints (e.g., a tiny RAM combined with a NAND Flash persistent storage badly adapted to random fine-grain updates). To tackle this challenge, we introduce three design principles, namely Write-Once Partitioning, Linear Pipelining and Background Linear Merging, and show how they can be combined to produce an embedded search engine reconciling high insert/delete/update rate and query scalability. We have implemented our search engine on a development board having a hardware configuration representative for smart objects. The experimental results demonstrate the scalability of the approach and its superiority compared to state of the art methods [28]. This work is part of Saliha Lallali's Ph.D. thesis.

Flash-based spatio-temporal indexing. The convergence of mobile computing, wireless communications and sensors has raised the development of many applications exploiting a massive flow of spatio-temporal data such as location-based services, participatory sensing, or traffic management [15]. Among the most active research topics in this area is the spatio-temporal data indexing. Nevertheless, since a few years a new fundamental parameter has made its entry on the database scene: the NAND flash storage. However, the peculiar characteristics of flash memory require redesigning the existing data storage and indexing techniques that were devised for magnetic hard-disks. In this study we propose TRIFL, an efficient and generic TRajectory Index for FLash. TRIFL is designed around the key requirements of trajectory indexing and flash storage. TRIFL is generic in the sense that it is efficient for both simple flash storage devices such as the SD cards and more powerful devices such as the solid state drives. In addition, TRIFL is supplied with an online self-tuning algorithm that allows adapting the index structure to the workload and the technical specifications of the flash storage device to maximize the index performance. Moreover, TRIFL achieves good performance with relatively low memory requirements, which makes the index appropriate for many application scenarios. The experimental evaluation shows that TRIFL outperforms the representative indexing methods on magnetic disks and flash disks. This work is part of Dai-Hai Ton That Ph.D. thesis, co-supervised by Iulian Sandu Popa.

6.2. Secure Global Computing on Asymmetric Architecture

Participants: Benjamin Nguyen [correspondent], Philippe Pucheral, Quoc-Cuong To.

Current applications, from complex sensor systems (e.g. quantified self) to online e-markets acquire vast quantities of personal information which usually ends-up on central servers. Decentralized architectures, devised to help individuals keep full control of their data, hinder global treatments and queries, impeding the development of services of great interest. In this study, we promote the idea of pushing the security to the edges of applications, through the use of secure hardware devices controlling the data at the place of their acquisition. To solve this problem, we propose secure distributed querying protocols based on the use of a tangible physical element of trust, reestablishing the capacity to perform global computations without revealing any sensitive information to central servers. This leads to execute global treatments on an asymmetric architecture, composed of a powerful, available and untrusted computing infrastructure (server or cloud), and a large set of low powered, highly disconnected trusted devices. Given our large scale data centric applications (e.g. nationwide surveys), we discard solutions based on secure multi-party computation, which do not scale. We have studied two different computing paradigms on this architecture: our first contribution was to study the execution of Privacy Preserving Data Publishing (PPDP) algorithms on such an architecture, and provided generic protocols to deal with all kinds of PPDP algorithms, which are robust against honest-but-curious and malicious adversaries [12], including vulgarization aspects [25]. Our second contribution was to study general SQL queries in this same execution context. For now, we have concentrated on the subset of SQL queries without joins, but including Group By and aggregates, and show how to secure their execution in the presence of honest-but-curious attackers [19]. Cost models and experiments demonstrate that this approach can scale to nationwide infrastructures [20][16]. This work is part of Quoc-Cuong To's Ph.D. thesis started in sept. 2012, and should be extended in particular to cover joins. We also plan to extend this general framework through a collaboration with INSA Centre Val de Loire, LIFO Lab and University of Paris Nord, LIPN lab, to study the secure execution of Map/Reduce on the Asymmetric Architecture.

6.3. Personal Cloud

Participants: Nicolas Anciaux [correspondent], Luc Bouganim, Athanasia Katsouraki, Benjamin Nguyen, Philippe Pucheral, Iulian Sandu Popa, Paul Tran Van.

We are witnessing an exponential increase in the acquisition of personal data about the individuals or produced by them. Today, this information is managed using Web applications, centralizing this data in cloud data servers, under the control of few Web majors [5]. However, it has now become clear that (1) centralizing millions of personal records exposes the data to very sophisticated attacks, linked to a very high potential benefit in case of success (millions of records being revealed), and (2) delegating the management of personal records without any tangible guarantee for the individuals leads to privacy violations, the data being potentially made accessible to other organizations (e.g., governments, commercial partners) and being subject to lucrative secondary usages (not advertised to the individuals). To face this situation, many recent initiatives push towards the emergence of the Personal Cloud paradigm. A personal cloud can be viewed as a personal server, owned by a given individual, which gives to its owner the ability to store her complete digital environment, synchronize it among various devices and share it with other individuals and applications under control. Many projects and startups currently investigate this solution, like OpenPDS, CozyCloud, OwnCloud, etc. In the SMIS team, we claim the need of a Secure Personal Cloud, and promote the introduction of a secure (tamper resistant) data engine in the architecture [11]. On this basis, we investigate new data sharing and dissemination models, where usage and access control rules endorsed by the individuals could be enforced. In 2014, we have presented this vision at EDBT'14 [18]. Several underlying research problems and perspectives have been presented in [11]. We have started a cooperation with the startup CozyCloud at the end of 2014. A contract was signed at the end of 2014 to integrate PlugDB in a CozyCloud instance and the PhD of Paul Tran Van (CIFRE SMIS-CozyCloud) has just started to explore new data sharing techniques which could be enforced in the secure personal cloud model. Athanasia Katsouraki is working on privacy issues and on adoption of the secure data engine in cooperation with the economists (CERDI) in the context of the Digital Society Institute (DSI).

6.4. Folk-IS

Participants: Nicolas Anciaux [correspondent], Luc Bouganim, Philippe Pucheral.

According to many studies, IT should become a key facilitator in establishing primary education, reducing mortality or supporting commercial initiatives in Least Developed Countries. The main barrier to the development of IT services in these regions is not only the lack of communication facilities, but also the lack of consistent information systems, security procedures, economic and legal support, as well as political commitment. In [5], we proposed the vision of trusted cells, a data platform for personal data services where the shared infrastructure (typically the cloud) is untrusted, while personal devices (such as smart phones, tablets or set-top box) are trusted execution environments. We revisited this vision to the context of LDCs. We proposed a new paradigm, that we call Folk-enabled Information System (Folk-IS), based on a fully decentralized and participatory approach, where each individual implements a small subset of a complete information system without the need for a shared networked infrastructure. As trusted cells, Folk-IS builds upon the emergence of highly secure, portable and low-cost storage and computing devices, called hereafter Smart Tokens. Here however, the focus is on low-cost of ownership, deployment and maintenance, and on the absence of a networked infrastructure. With Folk-IS and thanks to their smart tokens, people will transparently and opportunistically perform data management and networking tasks as they physically move, so that IT services are truly delivered by the crowd. We have published in [17] the Folk-IS vision and main principles, and in [13] a more detailed paper including technical challenges, specific to that approach and an exploitation and feasibility analysis of the Folk-IS vision.

SOCRATE Project-Team

6. New Results

6.1. Highlights of the Year

6.1.1. FIT/CortexLab Inauguration

FIT(Future Internet of Things) is a french Equipex (Équipement d'excellence) which aims to develop an experimental facility, a federated and competitive infrastructure with international visibility and a broad panel of customers. FIT is be composed of four main parts: a Network Operations Center (NOC), a set of Embedded Communicating Object (ECO) test-beds, a set of wireless OneLab test-beds, and a cognitive radio test-bed (CorteXlab) deployed by the Socrate team in the Citi lab. In 2014 the construction of the room was finished see Figure 5 . SDR nodes have installed in the room, 42 industrial PCs (Aplus Nuvo-3000E/P), 22 NI radio boards (USRP) and 18 Nutaq boards (PicoSDR, 2x2 and 4X4) can be programmed from internet now.

A very successfully inauguration took place on the 28th October 2014 [0], with the noticable venue of Vincent Poor, Dean of School of Engineering and Applied Science of Princeton University.

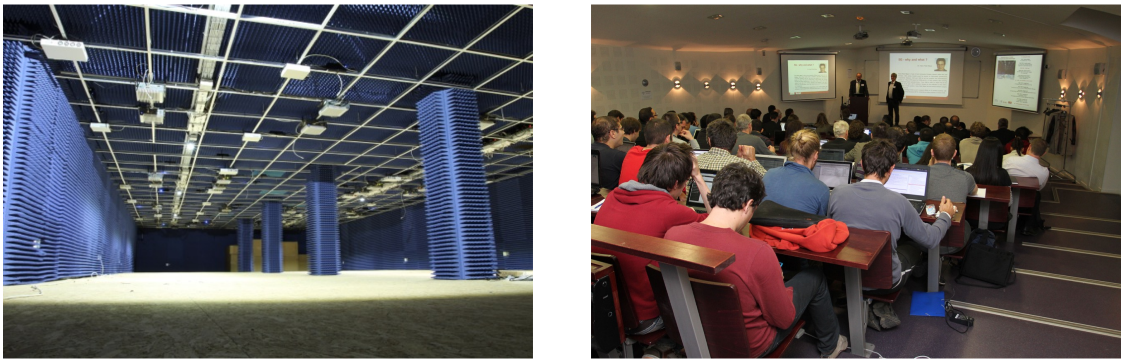


Figure 5. Photo of the FIT/CortexLab experimentation room installed and a snaptshot of the inauguration meeting

6.2. Flexible Radio Front-End

The innovative Wake-Up radio architecture proposed by the Socrate team, based on a classical WiFi standard with a specific OFDM pattern, has been deeply studied in theory and simulations [1], [25], [24]. Great enhancements on the sensitivity study, the choice of identifiers and the comparison of the energy consumption relative to classical systems have led to the development of a first prototype (ongoing work).

6.2.1. Wake-Up Radios

The innovative Wake-Up radio architecture proposed by the Socrate team, based on a classical WiFi standard with a specific OFDM pattern, has been deeply studied in theory and simulations [HUTU-JWCN][KHOUMERI-ECUMICT][HUTU-RWS]. Great enhancements on the sensitivity study, the choice of identifiers and the comparison of the energy consumption relative to classical systems have led to the development of a first prototype (ongoing work).

[0] http://www.inria.fr/centre/grenoble/actualites/inauguration-reussie-de-la-plateforme-cortexlab-equipex-fit

6.2.2. Full-Duplex systems

In the development of wideband OFDM Full-Duplex systems, [33] proposes an analysis of the impact of the thermal noise on the quality of the self-interference cancellation in such systems. A method is proposed to reduce the impact on the bit-error-rate by increasing the level of certain parts of the preamble in each frame. [35] add to the analog RF cancellation proposed previously a stage of digital cancellation enabling to increase more the performance of Full-Duplex terminals.

Furthermore, [34] extend the study to a dualband Full-Duplex systems, enabling the very promising combination of Full-Duplex and carrier aggregation. The proposed structure being sensitive to IQ impairments, a digital mitigation algorithm is also designed.

6.2.3. SDR for SRD

In collaboration with Orange labs, [32] analyses the requirements of an SDR gateway for urban networks collecting SRD (short range devices) information. This study is particularly focused on the ADC resolution, showing that the required resolution in realistic scenarios is too high, therefore emphasizing the need to develop specific hardware techniques.

6.2.4. Experimental Facilities

For the development of the CorteXlab testbed, lots of radio hardware and propagation constraints had to be taken into account [15], [14]. Moreover,[36] had proposed a first implementation of Full-Duplex on USRPs which is expected to be deployed on this tesbed.

Another testbed dedicated to the measurement of the energy consumption of radio devices was also designed and implemented.

6.3. Agile Radio Resource Sharing

This axis addresses the challenges relative to the network perspective of software radio. While the two other axes work on the design of the software radio nodes, we focus herein on their coexistence in a multi-user communications perspective. We are first interested in theoretical limits of some reference scenarios where trade-offs between spectral efficiency, energy efficiency, stability and/or fairness are analyzed. Our research activities are further driven by applicative frameworks. We focused on radio access networks with new results on energy efficiency-spectral efficiency trade-off in LTE networks and multi-band CSMA strategies in Wifi networks. We also studied pure random access and success probabilities for the challenging ultra-narrow band (UNB) technology of SigFox. Lot of efforts has been put on body area networks [8] with deep studies on positioning strategies and distributed decisions and information gathering. As mentioned above, our research follows three objectives:

- Establishing theoretical limits of cooperative wireless networks in the network information theory framework.
- Designing MAC procedures, coding and signal processing techniques for optimal transmissions (e.g. interference alignment).
- Developing distributed mechanisms for distributed decision at layer 1 and 2, using game theory, consensus and graph modeling.

6.3.1. Theoretical limits from information theory

The group strengthened his activities from a formal perspective in the framework of network information theory as initiated with the recruitment of Samir Perlaza and the sabbatical of Jean-Marie Gorce at Princeton University in the group of Prof. H. Vincent Poor. The first scenario is devoted to cellular networks with a random distribution on base stations. The main contribution concerns the broadcast channel (BC) generalized to a continuum of users. The second scenario concerns the interference channel (IC) and the main contribution is relative to the characterization of the Nash stable region for the interference channel with noisy feedback.

6.3.1.1. Broadcast channel with a continuum of users in a typical cell

The theoretical Energy efficiency-Spectral efficiency Pareto optimal front in a typical cell has been evaluated by associating stochastic geometry (Poisson point processes, PPP) and information theory.[21], the broadcast channel is extended to a continuum of users. We derived the theoretical uniform achievable rate with superposition coding principles. We show the potential gain of superposition coding techniques compared to the conventional time sharing. These results are however limited to Gaussian channels and the extension to the vector Gaussian channel is still under investigation. The PPP modeling for multi-cells has been also introduced as well as the price of interference management.

6.3.1.2. Interference Channel with feedback

The decentralized interference channel (DIC) with noisy feedback has been analyzed. In [31], all the rate-pairs that are achievable at a Nash equilibrium (NE) in the two-user linear deterministic symmetric decentralized interference channel (LD-S-DIC) with noisy feedback are identified. A second result provides closed form expressions for the PoA, which allows the full characterization of the reduction of the sum rate due to the anarchic behavior of all transmitter-receiver pairs. The price of anarchy (PoA) and the price of stability (PoS) of the game in which transmit-receiver pairs seek an optimal individual transmission rate are fully characterized in [9]. In particular, it is shown that in all interference regimes, there always exists at least one Pareto optimal Nash equilibrium (NE).

6.3.2. Coding, signal processing and MAC procedures for optimal transmissions

6.3.2.1. Implementation

While theoretical studies provide interesting insights about potential gain and limits of cognitive networks, the achievable efficiency may depend on practical issues related to quantization, synchronization and real-time processing limits. We developed the CortexLab facility offering a reproducible environment for fostering the validation of cooperative communication schemes. The first demo has been presented at the Infocom conference [28] and also at the Melbourne Greentouch meeting. We also contributed to the implementation and analysis of a cognitive transceiver for opportunistic networks [ref Maso JWCN]. The work first focused on a previously introduced dynamic spectrum access (DSA) - cognitive radio (CR) solution for primary-secondary coexistence in opportunistic orthogonal frequency division multiplexing (OFDM) networks, called cognitive interference alignment (CIA). The implementation is based on software-defined radio (SDR) and uses GNU Radio and the universal software radio peripheral (USRP) as the implementation toolkit. The proposed flexible transceiver architecture allows efficient on-the-fly reconfigurations of the physical layer into OFDM, CIA or a combination of both.

6.3.2.2. Interference alignment

In the framework of Greentouch, we studied interference alignment as a mean for improving the EE-SE tradeoff in cellular networks [43]. We combined theoretical studies with stochastic geometry and simulations to show the potential interest. We are also developing a demo with CorteXlab enhancing the IA capability from a real perspective.

6.3.2.3. Multiband MAC

In collaboration with CEA-Leti, we studied MAC strategies for multiband systems. The main idea is based on exploiting the multiband system as a slotted Aloha channel for the RTS/CTS initiation but keeping the total band as a whole for data transmission. We proved that this strategy outperforms classical approaches [39], [40], [30].

6.3.2.4. MAC for localization

In the context of the ANR Cormoran project, we account for radiolocation experiments aiming at both indoor navigation and mobility detection applications for Wireless Body Area Networks (WBAN) [7]. We also studied the relation between the MAC protocol and ranging techniques for localization. The impact of mobility on the distance estimation between 2 nodes of a Wireless Body Area Network (WBAN) by comparing the Two-Way Ranging (2WR) and Three-Way Ranging (3WR) protocols has been proposed in [23]. We also investigated the impact of mobility on the Motion Capture applications [22].

6.3.2.5. random access in Ultra-narrow band networks

Ultra narrow band (UNB) transmission is a very promising technology for low-throughput wireless sensor networks. This technology has already been deployed and has proved to be ultra-efficient for point-to-point communications in terms of power-efficiency, and coverage area. We studied the scalability of UNB for a multi-point to point network. In particular, we proposed a new multiple access scheme: random frequency division multiple access (R-FDMA) and studied the impact of the induced interference on the system performance in terms of bit error rate and outage probability [20]. We also analyzed the system performance in terms of bit error rate and outage probability [37].

6.3.3. Distributed decision mechanisms

Distributed decisions appear in many situations in the wireless world. Resource allocation, power management or relaying techniques are all expecting distributed decisions. To avoid strong coordination, distributed mechanisms inspired e.g. by game theory or consensus algorithms are appealing. Some of the results obtained below also rely on information theory but with a more important focus on algorithms and decision processes when several pairs of wireless transceivers are willing to simultaneously transmit in the same environment.

6.3.3.1. Cognitive radio networks

The problem of joint channel selection and power control is analyzed in the context of multiple-channel clustered ad-hoc networks in[ref Rose [3], i.e., decentralized networks in which radio devices are arranged into groups (clusters) and each cluster is managed by a central controller (CC). The problem is modeled by a game in normal form in which the corresponding utility functions are designed for making some of the Nash equilibria (NE) to coincide with the solutions to a global network optimization problem. A second scenario has been considered where multiple source-destination pairs communicate with each other via an energy harvesting relay [5]. The focus was put on the relay's strategies to distribute the harvested energy among the multiple users and their impact on the system performance. Specifically, a non-cooperative strategy that uses the energy harvested from the i-th source as the relay transmission power to the i-th destination is considered first. An auction based power allocation scheme is also proposed to achieve a better tradeoff between system performance and complexity.

6.3.3.2. Distributed decisions and consensus in MANETs

In the large research area of wireless body area networks, cooperative applications involving several users is attracting strong interests. This cooperation may target a simple information exchange or even some cooperative decision such as swarm coordination. We considered in [26]such a swarm of users moving in a common direction and we are interested in the mechanisms allowing to propagate and share some common information. We extend and improve a previous algorithm derived as a max-consensus approach. We describe a complete experimental setup deployed during a real bike race with 200 runners.

6.4. Software Radio Programming Model

6.4.1. Data Flow Programming

Software defined radio (SDR) technology has evolved rapidly and is now reaching market maturity. However, no standard has emerged for programming the new type of machine that will manage the access to the radio channel. Mickaël Dardaillon, Kevin Marquet, Tanguy Risset have been working in collaboration with the CEA LETI on compiling waveform for heterogeneous Multi-processor SoCs . This research leaded to a prototype compiler for the Magali MP-SoC developped in Mickael Dardaillon' PhD thesis (passed in November 2014) which was the first attempt to compile the SPDF format to a real architecture [18], [16], [17]. This study highlighted in particular the fact that SPDF was a good computation model for waveform description langage, easier to compile than dynamic dataflow format.

6.4.2. Non-volatile memory management for ultra low power systems

To enable non-trivial computation on very resource-constrained platforms powered by energy harvested from RF communications, an embedded OS has to save and restore program state to and from non-volatile memory. By doing so, the application program does not lose all progress when power is lost, which happens very often in environmentally-powered systems. This can be achieved [13] thanks to an incremental checkpointing scheme which aims at minimizing the amount of data written to non-volatile memory, while keeping the execution overhead as low as possible.

6.4.3. FPGA-based Implementation of physical Layers for SDR

A VHDL implementation of the three available options of the IEEE 802.15.4 physical layer was developed [29] in the context of FIT/CorteXlab. This parametrized design was validated on a Nutaq platform which combines Xilinx Virtex-6 FPGA and tunable Radio420x RF transceiver. This work participates to the building of an open source hardware SDR library similar to GNU radio but targeted to FPGA-based platforms.

6.4.4. Towards filters and functions computing just right

A FIR filter is specified by its coefficients (real numbers) and its input and output formats. The implementation of a FIR should be as accurate as its output format allows, but no more. This very simple specification enables the automatic construction of FIR filter implementations that are provably accurate at a minimal hardware cost [19]. The corresponding FIR generator is available in FloPoCo.

The fixed-point Atan2 function is very useful to recover the phase of a complex signal. A careful study of three implementation techniques (including a novel one based on two-variable quadratic approximation) shows that, on current FPGAs, the good old CORDIC technique is more efficient than multiplier-based techniques [46].

SPADES Team

6. New Results

6.1. Components and Contracts

Participant: Jean-Bernard Stefani.

6.1.1. Location graph model

The design of configurable systems can be streamlined and made more systematic by adopting a component-based structure, as demonstrated with the Fractal component model [2]. However, the formal foundations for configurable component-based systems, featuring higher-order capabilities where components can be dynamically instantiated and passivated, and non-hierarchical structures where components can be contained in different composites at the same time, are still an open topic. We have developed recently the location graph model [15], where components are understood as graphs of locations hosting higher-order processes, and where component structures can be arbitrary graphs. We have developed a compositional operational semantics for the location graph model, which is parametric with respect to the family of processes. We have shown that the location graph model constitutes a conservative extension of a previous model, called CAB, that captures the key features of the BIP component model [5]. We have further worked on the behavioral theory of the location graph model, characterizing contextual equivalence in the model by means of a higher-order bisimularity relation, and begun the study of the encoding of different models, including the Synchronized Hyperedge Replacement model [45].

6.2. Real-Time multicore programming

Participants: Vagelis Bebelis, Adnan Bouakaz, Pascal Fradet, Alain Girault, Gregor Goessler, Jean-Bernard Stefani, Sophie Quinton, Partha Roop, Eugene Yip.

6.2.1. Analysis and scheduling of parametric dataflow models

Recent data-flow programming environments support applications whose behavior is characterized by dynamic variations in resource requirements. The high expressive power of the underlying models (*e.g.*, Kahn Process Networks or the CAL actor language) makes it challenging to ensure predictable behavior. In particular, checking *liveness* (*i.e.*, no part of the system will deadlock) and *boundedness* (*i.e.*, the system can be executed in finite memory) is known to be hard or even undecidable for such models. This situation is troublesome for the design of high-quality embedded systems.

Recently, we have introduced the *schedulable parametric data-flow* (SPDF) MoC for dynamic streaming applications [47]. SPDF extends the standard dataflow model by allowing rates to be parametric. Last year, we have proposed the *Boolean Parametric Data Flow* (BPDF) MoC which combines integer parameters (to express dynamic rates) and boolean parameters (to express the activation and deactivation of communication channels). High dynamicity is provided by integer parameters which can change at each basic iteration and boolean parameters which can change even within the iteration. We have presented static analyses which ensure the liveness and the boundedness of BDPF graphs.

Recently, we have proposed a generic and flexible framework to generate parallel schedules for BPDF applications [16]. The parametric dataflow graph is associated with user-defined specific constraints aimed at minimizing, timing, buffer sizes, power consumption, or other criteria. The scheduling algorithm executes with minimal overhead and can be adapted to different scheduling policies just by changing some constraints. The safety of both the dataflow graph and constraints can be checked statically and all schedules are guaranteed to be bounded and deadlock free. Our case studies are video decoders for high definition video streaming such as VC-1. One of the target architectures is the STHORM many-core platform designed by STMicroelectronics.

This research is the central topic of Vagelis Bebelis' PhD thesis. It is conducted in collaboration with STMicroelectronics.

6.2.2. Typical Worst-Case Analysis of real-time systems

Weakly hard time constraints have been proposed for applications where occasional deadline misses are permitted. We have recently developed Typical Worst Case Analysis (TWCA) to exploit similar constraints and bound response times of systems with sporadic overload. This year, we have applied this approach to a real-life automotive network [14]. Additionally, we have extended the approach for static priority preemptive (SPP) and static priority non-preemptive (SPNP) scheduling to determine the maximum number of deadline misses of a given task [21]. The approach is based on an optimization problem which trades off higher priority interference versus miss count. We formally derived a lattice structure for these combinations that lays the ground for an integer linear programming (ILP) formulation. The ILP solution was evaluated and provided far better results than previous TWCA.

In parallel, we have contributed to a systematic co-engineering approach that integrates TWCA into functional analysis [19]. We combine physical, control and timing models by representing them as a network of hybrid automata. Closed-loop properties can then be verified on this hybrid automata network by using standard model checkers for hybrid systems. The use of the Logical Execution Time (LET) semantics where data is written back deterministically at the typical worst-case response time (rather than the usual worst-case bound) is a new and particularly powerful approach for addressing the computational complexity of the model checking problem.

6.2.3. Time predictable programming

In the context of the RIPPES associated team with UC Berkeley and U Auckland, we have finalized ongoing work on our synchronous programming language for time predictability PRET-C [10]. PRET-C extends C with synchronous constructs inspired by ESTEREL, to allow an easy programming of concurrent reactive programs. These constructs allow the programmer to express concurrency, interaction with the environment, looping, and a synchronization barrier (like the `pause` statement in ESTEREL). PRET-C's semantics is deterministic, and it can be efficiently compiled towards sequential code, either executed on a dedicated processor for the best predictability of the program's Worst-Case Reaction Time (WCRT), or executed on a generic processor.

We have also continued our work on FOREC, a time predictable synchronous programming language for multi-core chips. Like PRET-C, it extends C with a small set of ESTEREL-like synchronous primitives. FOREC threads communicate with each other via shared variables, the values of which are combined at the end of each tick to maintain deterministic execution. FOREC is compiled into threads that are then statically scheduled for a target multi-core chip. This is the main difference with PRET-C. We have finalized the semantics of FOREC, which led us to propose several ways to combine shared variables at the tick boundaries, such that the semantics remains deterministic. This part was inspired by the so-called concurrent revisions [38].

Finally, with colleagues from the former ARTISTDESIGN European Network of Excellence, we have also participated in a survey on predictable embedded systems [11].

6.2.4. Tradeoff exploration between energy consumption and execution time

We have continued our work on multi-criteria scheduling, in the particular context of dynamic applications that are launched and terminated on an embedded multi-core chip, under execution time and energy consumption constraints. We have proposed a two layer adaptive scheduling method. In the first layer, each application (represented as a DAG of tasks) is scheduled statically on sets of cores: 2 cores, 3 cores, 4 cores, and so on. For each size of these sets (2, 3, 4, ...), there may be only one topology or several topologies. For instance, for 2 or 3 cores there is only one topology (a "line"), while for 4 cores there are three distinct topologies ("line", "square", and "T shape"). Moreover, for each topology, we generate statically several schedules, each one subject to a different total energy consumption constraint, and consequently with a different Worst-Case Reaction Time (WCRT). Coping with the energy consumption constraints is achieved thanks to Dynamic Frequency and Voltage Scaling (DVFS). In the second layer, we use these pre-generated static schedules to reconfigure dynamically the applications running on the multi-core each time a new application is launched or an existing one is stopped. The goal of the second layer is to perform a global optimization of the configuration, such that each running application meets a pre-defined quality-of-service constraint (translated into an upper bound on its WCRT) and such that the total energy consumption is minimized. For this, we (1) allocate a sufficient number of cores to each active application, (2) allocate the unassigned cores to the applications yielding the largest gain in energy, and (3) choose for each application the best topology for its subset of cores (i.e., better than the by default "line" topology).

This is a joint work with Ismail Assayad (U. Casablanca, Morocco) who visits the team regularly.

6.3. Language Based Fault-Tolerance

Participants: Dmitry Burlyaev, Pascal Fradet, Alain Girault, Yoann Geoffroy, Gregor Goessler, Jean-Bernard Stefani.

6.3.1. Automatic transformations for fault tolerant circuits

In the past years, we have studied the implementation of specific fault tolerance techniques in real-time embedded systems using program transformation [1]. We are now investigating the use of automatic transformations to ensure fault-tolerance properties in digital circuits. To this aim, we consider program transformations for hardware description languages (HDL). We consider both single-event upsets (SEU) and single-event transients (SET) and fault models of the form "at most 1 SEU or SET within n clock signals".

We have expressed several variants of triple modular redundancy (TMR) as program transformations. We have proposed a verification-based approach to minimize the number of voters in TMR [17]. Our technique guarantees that the resulting circuit *(i)* is fault tolerant to the soft-errors defined by the fault model and *(ii)* is functionally equivalent to the initial one. Our approach operates at the logic level and takes into account the input and output interface specifications of the circuit. Its implementation makes use of graph traversal algorithms, fixed-point iterations, and BDDs. Experimental results on the ITC'99 benchmark suite indicate that our method significantly decreases the number of inserted voters, which entails a hardware reduction of up to 55% and a clock frequency increase of up to 35% compared to full TMR. We address scalability issues arising from formal verification with approximations and assess their efficiency and precision.

We have proposed novel fault-tolerance transformations based on time-redundancy. In particular, we have presented a transformation using double-time redundancy (DTR) coupled with micro-checkpointing, rollback and a speedup mode [18]. The approach is capable to mask any SET every 10 cycles and keeps the same input/output behavior regardless error occurrences. Experimental results on the ITC'99 benchmark suite indicate that the hardware overhead is 2.7 to 6.1 times smaller than full TMR with double loss in throughput. It is an interesting alternative to TMR for logic intensive designs.

We have also designed a transformation that allows the circuit to change its level of time-redundancy. This feature permits to dynamically and temporarily give up (resp. increase) fault-tolerance and speed up (resp. slow down) the circuit. The motivations for such changes can be based on the observed change in radiation environment or the processing of (non)critical data. These different time redundancy transformations have been patented [23]

We have started the formal certification of such transformations using the Coq proof assistant [40]. The transformations are described on a simple gate-level hardware description language inspired from μFP [68]. The fault-model is described in the operational semantics of the language. The main theorem states that, for any circuit, for any input stream and for any SET allowed by the fault-model, its transformed version produces a correct output. A TMR and triple time redundancy transformations have already been proved correct. The proof of the DTR transformation is in progress.

6.3.2. Concurrent flexible reversibility

In the recent years, we have been investigating reversible concurrent computation, and investigated various reversible concurrent programming models, with the hope that reversibility can shed some light on the common semantic features underlying various forms of fault recovery techniques (including, exceptions, transactions, and checkpoint/rollback schemes).

We have revisited our encoding of our reversible higher-order π-calculus in (a variant of) the higher-order π-calculus, in order to obtain a much tighter result than our original encoding. In essence, we now have a form of strong bisimilarity (modulo administrative reductions) between a reversible higher-order π-calculus process and its translation in higher-order π. We have also studied the relation between the causality information used in our reversible higher-order π and a causal higher-order π-calculus, inspired by the causal π-calculus [35]. This work has been submitted for publication [24]. This work was done in collaboration with Inria teams FOCUS in Bologna, as part of the ANR REVER project.

6.3.3. Blaming in component-based systems

The failure of one component may entail a cascade of failures in other components; several components may also fail independently. In such cases, elucidating the exact scenario that led to the failure is a complex and tedious task that requires significant expertise.

The notion of causality *(did an event e cause an event e'?)* has been studied in many disciplines, including philosophy, logic, statistics, and law. The definitions of causality studied in these disciplines usually amount to variants of the counterfactual test "e is a cause of e' if both e and e' have occurred, and in a world that is as close as possible to the actual world but where e does not occur, e' does not occur either". Surprisingly, the study of logical causality has so far received little attention in computer science, with the notable exception of [51] and its instantiations. However, this approach relies on a causal model that may not be known, for instance in presence of black-box components. For such systems, we have been developing a framework for blaming that helps us establish the causal relationship between component failures and system failures, given an observed system execution trace. The analysis is based on a formalization of counterfactual reasoning. We have shown in [12] how our approach can be used for log analysis to help establishing liability in the context of legal contracts.

We have proposed in [6] an approach for blaming in component-based real-time systems whose component specifications are given as timed automata. The analysis is based on a single execution trace violating a safety property P. We have formalized blaming using counterfactual reasoning to distinguish component failures that actually contributed to the outcome from failures that had no impact on the violation of P. We have shown how to effectively implement blaming by reducing it to a model-checking problem for timed automata. The approach has been implemented in LOCA (Section 5.1.1). We have further demonstrated the feasibility of our approach on the model of a dual-chamber implantable pacemaker.

6.3.4. Synthesis and implementation of fault-tolerant embedded systems

We have integrated a complete workflow to synthesize and implement correct-by-construction fault tolerant distributed embedded systems consisting of real-time periodic tasks. Correct-by-construction is provided by the use of discrete controller synthesis [63] (DCS), a formal method thanks to which we are able to guarantee that the synthe-sized controlled system satisfies the functionality of its tasks even in the presence of processor failures. For this step, our workflow uses the Heptagon domain specific language [43] and the Sigali DCS tool [59]. The correct implementation of the resulting distributed system is a challenge, all the more since the controller itself must be tolerant to the processor failures. We achieve this step thanks to the libDGALS real-time library [22] (1) to generate the glue code that will migrate the tasks upon processor failures, maintaining their internal state through migration, and (2) to make the synthesized controller itself fault-tolerant. We have demonstrated the feasibility of our work-flow on a multi-tasks multi-processor fault-tolerant distributed system.

SPECFUN Project-Team

6. New Results

6.1. Highlights of the Year

Two results are particularly important this year, our computer-checked proof [11] of irrationality of $\zeta(3)$ and our new algorithm [19] for the integration of multiple integrals. The former is our first success in the merger between computer algebra and formal methods, and stimulates further research in this direction around special functions and creative telescoping. The latter has made a large class of integrals possible in practice, thus allowing us to compute a challenging list of integrals related to famous Calabi–Yau various; it has also received attention by physicists.

6.2. A formal proof of the irrationality of $\zeta(3)$

We have obtained a formal proof, machine-checked by the Coq proof assistant, of the irrationality of the constant $\zeta(3)$, that is, the evaluation at 3 of the Riemann zeta function of number theory. The result has been known in mathematics since the French mathematician Apéry's work in 1978, and several alternative proofs have been given since then. Our formalized result is the first complete proof by the computer (under the single assumption of the asymptotic behavior of the least common multiple of the first n natural numbers). The core of this formal proof is based on (untrusted) computer-algebra calculations performed outside the proof assistant with the Mgfun Maple library developed by members of the team in the past. Then, we verify formally and a posteriori the desired properties of the objects computed by Maple and complete the proof of irrationality. This work [11] was formally presented at the conference on interactive theorem proving, ITP'14, and also as talks at MSC 2014 (Mathematical Structures of Computation) [0], at the meeting MAP 2014 of the community on mathematics, algorithms and proofs [0], and at JNCF'14, the meeting of the French computer-algebra community [0].

6.3. Criterion for the existence of telescopers for mixed hypergeometric terms

Creative telescoping is a process that determines a univariate recurrence satisfied by the sum of a summand described by a system of bivariate recurrences. For hypergeometric summands, that is, summands given by first-order linear recurrences, this has led to Zeilberger's algorithm in the early 1990s, since then followed by a large number of works, including a natural counterpart for integration. The history of creative-telescoping algorithms was surveyed this year in Chyzak's HDR [1]. Also this year, we presented in [6] a criterion for the existence of telescopers for mixed hypergeometric terms, which is based on additive and multiplicative decompositions. The criterion had enabled us to determine the termination of Zeilberger's algorithms for mixed hypergeometric inputs prior to any costly computations, and to verify that certain indefinite sums do not satisfy any polynomial differential equation.

6.4. Integration of rational functions

Periods of rational integrals are specific integrals, with respect to one or several variables, whose integrand is a rational function and whose domain of integration is closed. Periods with a parameter are classically known to satisfy linear differential equations of a type called Picard-Fuchs equations. As for other special-function manipulations, handling periods through those differential equations is a good way to actually compute them, and this was the topic of Pierre Lairez' PhD, defended this year [2].

[0] http://smc2014.univ-lyon1.fr/
[0] http://perso.crans.org/cohen/map2014/
[0] http://www.lifl.fr/jncf2014/

Computing multivariate integrals is one speciality of the team and our algorithms are known to treat much more general integrals than just periods of rational integrals. However, integration is still slow in practice when the number of variables goes increasing. By looking at periods of rational function, the hope is to obtain relevant complexity bounds and faster algorithms.

The goal of reaching relevant theoretical complexity bounds has been reached last year [35] but a practically fast algorithm was still missing. This year, we described a new algorithm which is efficient in practice [19], though its complexity is not known. This algorithm allows to compute quickly integrals that are too big to be computed with previous algorithms. As a challenging benchmark, we computed 210 integrals given by Batyrev and Kreuzer in their work on Calabi–Yau varieties. This achievement gave strong visibility to the paper and allowed a quick dissemination of the implementation, which is provided in Magma under a CeCILL B license. The algorithm is now used on a regular basis by several teams. We know of:

- Tom Coates' team (Dpt. of Mathematics, Imperial College, London, UK), which uses the software in their work about mirror symmetry and classification of Fano varieties;
- Duco van Straten (Institute of Mathematics, University of Mainz, Germany), who uses the software in his work in algebraic geometry;
- Gert Alkmvist (Dpt. of Mathematics, University of Lund, Sweden), who uses the software in his work of enumerating the Calabi–Yau differential equations.

6.5. Efficient algorithms for linear differential equations in positive characteristic

The p-curvature of a linear differential operator in characteristic p is a matrix that measures to what extent the space of polynomial solutions of the operator has dimension close to its order. This makes the p-curvature a useful tool in concrete applications, like in combinatorics and statistical physics, where it serves for instance as an a posteriori certification filter for differential operators obtained by guessing techniques. In [9], we designed a new algorithm for computing the characteristic polynomial of the p-curvature in sublinear time $\tilde{\mathrm{O}}(p^{0.5})$. Prior to this work, the fastest algorithms for this task, and even for the subtask of deciding nilpotency of the p-curvature, had had merely slightly subquadratic complexity $\tilde{\mathrm{O}}(p^{1.79})$. The new algorithm is also efficient in practice: it allows to test the nilpotency of the p-curvature for primes p of order 10^6, for which the p-curvature itself is impossible to compute using current algorithms.

6.6. Efficient algorithms for rational first integrals

We presented in [4] fast algorithms for computing rational first integrals with degree bounded by N of a planar polynomial vector field of degree $d \leq N$. The main novelty is that such rational first integrals are obtained by computing via systems of linear equations instead of systems of quadratic equations. This leads to a probabilistic algorithm with arithmetic complexity $\tilde{\mathrm{O}}(N^{2\omega})$ and to a deterministic algorithm for solving the problem in $\tilde{\mathrm{O}}(d^2 N^{2\omega+1})$ arithmetic operations, where ω is the exponent of linear algebra. By comparison, the best previous algorithm uses at least $d^{\omega+1} N^{4\omega+4}$ arithmetic operations. Our new algorithms are moreover very efficient in practice.

6.7. Computation of necessary integrability conditions for parametrized Hamiltonian systems

Let $V(\mathbf{q}_1, \mathbf{q}_2)$ be a homogeneous function whose coefficients depend rationally on parameters $\mathbf{a}_1, ..., \mathbf{a}_n$. In [10] we designed an algorithm to compute polynomial necessary conditions on the parameters $(\mathbf{a}_1, ..., \mathbf{a}_n)$ such that the dynamical system associated to the potential V is integrable. These conditions originate from those of the classical Morales-Ramis-Simó integrability criterion. The implementation of the algorithm allows to treat applications that were out of reach before, for instance concerning the non-integrability of polynomial potentials up to degree 9. Another striking application is the first complete proof of the non-integrability of the collinear three-body problem.

6.8. Non-D-finite excursions in the quarter plane

Counting lattice paths obeying various geometric constraints is a classical topic in combinatorics and probability theory. Many recent works deal with the enumeration of 2-dimensional walks with prescribed steps confined to the positive quadrant. A large part of the effort has been devoted to the classification of classes of walks according to the nature of equations that they satisfy (linear, polynomial, differential, etc). Equivalently, this provides properties of the classes of walks according to the algebraic nature of their enumerative series: whether rational, algebraic, D-finite, etc. The classification is now complete for walks with unit steps: the trivariate generating function of the numbers of walks with given length and prescribed ending point is D-finite if and only if a certain group associated with the step set is finite. We proved in [5] a refinement of this result: we showed that the sequence of numbers of excursions (finite paths starting and ending at the origin) in the quarter plane corresponding to a nonsingular step set with infinite group does not satisfy any nontrivial linear recurrence with polynomial coefficients. This solves an open problem in the field of lattice-path combinatorics.

6.9. A human proof of the Gessel conjecture

Gessel walks are planar walks confined to the positive quarter plane, that move by unit steps in any of the following directions: West, North-East, East and South-West. In 2001, Ira Gessel conjectured a closed-form expression for the number of Gessel walks of a given length starting and ending at the origin. In 2008, Kauers, Koutschan and Zeilberger gave a computer-aided proof of this conjecture. The same year, Bostan and Kauers showed, using again computer algebra tools, that the trivariate generating function of Gessel walks is algebraic. We propose in [17] the first “human proofs” of these results. They are derived from a new expression for the generating function of Gessel walks.

6.10. Enumeration of 3-dimensional lattice walks confined to the positive octant

We explored in [3] the classification problem for 3-dimensional walks with unit steps confined to the positive octant. The first difficulty is their number: there are 11 074 225 cases (instead of 79 in dimension 2). In our work, we focused on the 35 548 that have at most six steps. We applied to them a combined approach, first experimental and then rigorous. Among the 35 548 cases, we first found 170 cases with a finite group; in the remaining cases, our experiments suggest that the group is infinite. We then rigorously proved D-finiteness of the generating series in all the 170 cases, with the exception of 19 intriguing step sets for which the nature of the generating function still remains unclear. In two challenging cases, no human proof is currently known, and we derived computer-algebra proofs, thus constituting the first proofs for those two step sets.

6.11. Asymptotic expansions for linear homogeneous divide-and-conquer recurrences: Algebraic and analytic approaches collated

Linear divide-and-conquer recurrences are a classical topic in computer science, but they are often dealt with in an offhand way. Particularly the subtle oscillations they show are usually not emphasized. After having elaborated last year a new approach to the asymptotic study of such recurrences, we provide in [7] a comparison with an older approach based on number theoretic tools as Dirichlet series and residue computation. The most striking aspect of the linear approach is the simplicity and the ease of use. Reduction to normal Jordan form, computation of a joint spectral radius, dealing with a dilatation equation are all workable with a computer-algebra system. Moreover these concepts are better known by computer scientists than those of complex analysis and analytic number theory. So there is hope that this approach will more easily gain acceptance among computer scientists.

6.12. Asynchronous interaction with Coq

We have integrated the Coq proof assistant with the PIDE architecture [13], [12] ("prover integrated development environment"). The architecture is aimed at asynchronous, parallel interaction with proof assistants, originally aimed at the Isabelle proof assistant, and is tied in heavily with a plugin that allows the jEdit editor to work with proof assistants. We have made several generalizations to the PIDE architecture to accommodate for more provers than just Isabelle, and adapted Coq to understand the core protocol: this delivered a working system in about two man-months; further work improved the connection and added novel functionalities to the interface. The tool has also been presented informally at seminars at the University of Dundee and the Université Paris 13.

SPIRALS Team

6. New Results

6.1. Highlights of the Year

In 2014, we are proud to have organized the 17th ACM SIGSOFT International Conference on Component-Based Software Engineering and Software Architecture (CompArch) that has been held in Lille from 30 June to 3 July 2014.

CompArch is the main conference of the ACM SIGSOFT group on software architectures and software components. The conference is held alternatively in North America and in Europe. The 17th edition has been held this year in France for the first time. The conference brings together about 100 researchers from the academia and the industry.

6.2. Distributed Context Monitoring

In 2014, we obtained some new results in the area of distributed context monitoring solutions to support the development of self-optimising software systems. Context monitoring has emerged as a key capability in various domains to connect software systems to the underlying hardware platform or to the physical world (in the case of ubiquitous systems). In particular, we have investigated to the capability of inferring high-level contextual situations from a large volume of raw data collected from a single device or in the wild. Both hardware (*e.g.*, accelerometer) or software (*e.g.*, performance counters) sensors tend to continuously produce raw data that a context monitoring solution has to quickly filter, process, and convert it into information that can be used by an application or understood by a user.

As a result of the PhD thesis of Adel Noureddine [14], defended in March 2014, we have developed a middleware toolkit to support *in-depth context monitoring* in the domain of green computing. In particular, we introduce a software library, named POWERAPI, that can estimate the power consumption in real-time at various granularities of software: from system processes to code methods (see Section 5.3). This non-invasive solution provides accurate insights on energy hotspots of software and can be used to derive the energy profile of any software library, thus guiding the developers in optimising the energy consumption of their developments.

As a result of the PhD thesis of Nicolas Haderer [12], defended in November 2014, we have contributed to the development of a middleware platform to support *in-breadth context monitoring* in the area of mobile computing. In particular, we promote the distributed middleware solution APISENSE® as an efficient approach to deploy mobile crowd-sensing tasks across a large population of volunteer participants (see Section 5.1). In particular, APISENSE® includes a task orchestration algorithm that preserves the privacy and the battery of sensing devices, while maintaining specific sensing coverage objectives (including time and space dimensions). The server-side infrastructure of APISENSE® is generated from a dedicated software product line, while the implementation is based on the FRASCATI platform (see Section 5.2).

6.3. Design and Runtime Support for Cloud Computing

In 2014, we obtained some new results in the domain both of the design and the runtime support of distributed applications for multi-cloud systems. The purpose is to deal with applications that span across several different cloud systems. Several reasons justify such a goal. For example, in order to avoid the so-called vendor lock-in syndrome, cloud application stakeholders need to be able to migrate as easily as possible their assets from one cloud system to another one. Other examples include the possibility of introducing diversify and fault-tolerance by deploying applications on different cloud systems, or hot migrating applications where computing resources are less expensive.

For the design of multi-cloud systems, we proposed a solution based on software product lines (SPL) [90] and ontologies. In order to specify the variability of such environments, we extended SPL with attributes, cardinalities, and constraints. In order to enable the evolution of these environments, we provided an automated support for maintaining the consistency based on constraint programming. Finally, we proposed an ontology based approach to bridge the gap between the concepts and artefacts defined by different cloud systems. This global solution is the result of the PhD thesis of Clément Quinton [16] that was defended in October 2014, and has been partially supported by the FP7 PaaSage project (see Section 8.3).

For the runtime support of multi cloud systems, we proposed the SOCLOUD platform. This solution enables to deploy, execute and manage an application that spans on several different cloud systems. SOCLOUD tackles the challenges of portability, provisioning, elasticity, and high availability. SOCLOUD defines a component-based and service-oriented architecture that provides an unified view of a set of cloud systems. SOCLOUD is the result of the PhD thesis of Fawaz Paraiso [15] that was defended in June 2014. SOCLOUD is implemented on top of the FRASCATI platform (see Section 5.2).

6.4. Extraction and Analysis of Knowledge for Automatic Software Repair

Automated software repair aims at assisting developers in order to improve the quality of software systems, for example by recommending some repair actions to fix bugs. Matias Martinez has presented in his PhD thesis [13] that was defended in June 2014, new results in this domain. These results aim at reducing the search space when repairing a software system. The solution relies on two techniques. The first one consists in building change models learnt from repairs performed by other developers. These repairs are mined from existing software repositories of open source projects, and analysed based on their types and frequencies. The second proposed technique is based on the inherent redundancy of code patterns. The assumption is that the probability that the repair code for a particular kind of defect is already present in the software system under study is high. We then take advantage of this inherent redundancy to reduce the search space when looking for repair actions.

STARS Project-Team

6. New Results

6.1. Highlights of the Year

NeoSensys, a spin off of the Stars team which aims at commercializing video surveillance solutions for the retail domain, has been created in September 2014.

6.2. Introduction

This year Stars has proposed new algorithms related to its three main research axes : perception for activity recognition, semantic activity recognition and software engineering for activity recognition.

6.2.1. Perception for Activity Recognition

Participants: Julien Badie, Slawomir Bak, Piotr Bilinski, François Brémond, Bernard Boulay, Guillaume Charpiat, Duc Phu Chau, Etienne Corvée, Carolina Garate, Michal Koperski, Ratnesh Kumar, Filipe Martins, Malik Souded, Anh Tuan Nghiem, Sofia Zaidenberg, Monique Thonnat.

For perception, the main achievements are:

- Our new covariance descriptor has led to many publications and applications already. The work on this topic is now more about the precise use of the descriptor in varied applications than the design of new descriptors.
- The new action descriptors have led to finer gesture classification. As our target application is the detection of the Alzheimer syndrome from gesture analysis, which requires still finer descriptors, we will continue the work on this topic.
- The different shape priors developed (for shape growth enforcement, shape matching, articulated motion) have been formulated and designed so that efficient optimization tools could be used, leading to global optimality guarantees. These particular problems can thus be considered as solved, but there is still much work to be done on shape and related optimization, in particular to obtain shape statistics for human action recognition.
- The success obtained in the control of trackers is a proof of concept, but this work still needs to be pursued to get more practical and to be applied on more real world videos.

More precisely, the new results for perception for activity recognition are:

- People Detection for Crowded Scenes (6.3),
- Walking Speed Detection on a Treadmill using an RGB-D camera : experimentations and results (6.4),
- Head detection using RGB-D camera (6.5),
- Video Segmentation and Multiple Object Tracking (6.6),
- Enforcing Monotonous Shape Growth or Shrinkage in Video Segmentation (6.7),
- Multi-label Image Segmentation with Partition Trees and Shape Prior (6.8),
- Automatic Tracker Selection and Parameter Tuning for Multi-object Tracking (6.9),
- An Approach to Improve Multi-object Tracker Quality using Discriminative Appearances and Motion Model Descriptor (6.10),
- Person re-identification by pose priors(6.11),
- Global tracker : an online evaluation framework to improve tracking quality (6.12),
- Human action recognition in videos (6.13),
- Action Recognition using 3D Trajectories with Hierarchical Classifier (6.14),
- Action Recognition using Video Brownian Covariance Descriptor for Human (6.15),
- Towards Unsupervised Sudden Group Movement Discovery for Video Surveillance (6.16).

6.2.2. Semantic Activity Recognition

Participants: Vania Bogorny, Luis Campos Alvares, Vasanth Bathrinarayanan, Guillaume Charpiat, Duc Phu Chau, Serhan Cosar, Carlos F. Crispim Junior, Giuseppe Donatielo, Baptiste Fosty, Carolina Garate, Alvaro Gomez Uria Covella, Alexandra Konig, Farhood Negin, Anh-Tuan Nghiem, Philippe Robert, Carola Strumia.

For activity recognition, the main advances on challenging topics are:

- The utilization by clinicians for their everyday work of a first monitoring system able to recognize complex activities, to evaluate in real-time older people performance in an ecological room at Nice Hospital.
- The successful processing of over 80 older people videos and matching their performance for autonomy at home (e.g. walking efficiency) and cognitive disorders (e.g. realisations of executive tasks) with gold standard scales (e.g. NPI, MMSE). This research work contributes to the early detection of deteriorated health status and the early diagnosis of illness.
- The fusion of events coming from camera networks and heterogeneous sensors (e.g. RGB videos, Depth maps, audio, accelerometers).
- The management of the uncertainty of primitive events.
- The generation of event models in an unsupervised manner.

For this research axis, the contributions are :

- Autonomous Monitoring for Securing European Ports (6.17),
- Video Understanding for Group Behavior Analysis (6.18),
- Evaluation of an event detection framework for older people monitoring: from minute to hour-scale monitoring and Patients autonomy and dementia assessment (6.19),
- Uncertainty Modeling Framework for Constraint-based Event Detection in Vision Systems (6.20),
- Assisted Serious Game for older people (6.21),
- Enhancing Pre-defined Event Models using Unsupervised Learning (6.22),
- Using Dense Trajectories to Enhance Unsupervised Action Discovery (6.23),
- Abnormal Event Detection in Videos and Group Behavior Analysis (6.24).

6.2.3. Software Engineering for Activity Recognition

Participants: François Brémond, Daniel Gaffé, Sabine Moisan, Annie Ressouche, Jean-Paul Rigault, Omar Abdalla, Mohamed Bouatira, Ines Sarray, Luis-Emiliano Sanchez.

For the software engineering part, the main achievements are the Software Engineering methods and tools applied to video analysis. We have demonstrated that these approaches are appropriate and useful for video analysis systems:

- Run time adaptation using MDE is a promising approach. Our current prototype resorts to tools and technologies which were readily available. This made possible a proof of concepts.
- Introducing metrics in feature models was valuable to reduce the huge set of valid configurations after a dynamic context change and to provide a real time selection of an appropriate running configuration.
- The synchronous approach is well suited to describe reactive systems in a generic way, it has a well-established formal foundation allowing for automatic proofs, and it interfaces nicely with most model-checkers.

The contributions for this research axis are:

- Model-Driven Engineering for Activity Recognition Systems(6.25),
- Scenario Analysis Module (6.26),
- The Clem Workflow (6.27),
- Multiple Services for Device Adaptive Platform for Scenario Recognition (6.28).

6.3. People Detection for Crowded Scenes

Participants: Malik Souded, François Brémond.

keywords: people detection, crowded scenes, features, boosting.

This works aims at proposing an efficient people detection algorithm which can deal with crowded scenes.

6.3.1. Early Work

We have previously proposed an approach which optimizes state-of-the-art methods [Tuzel 2007, Yao 2008], based on training cascade of classifiers using LogitBoost algorithm on region covariance descriptors. This approach performs in real time and provides good detection performances in low to medium density scenes (see some examples in figure 10). However, this approach shows its limits on crowded scenes. Both detection accuracy and detection time are highly impacted in this case. The detection time increases dramatically due to the number of people in images, which forces the evaluation of many cascade levels, while the numerous partial occlusions highly decrease the detection rate (the considered detector is a full-body detector). To deal with these issues, we are working on a new approach.

6.3.2. Current Work

Our approach is based on training a cascade of classifiers using Boosting algorithms too, but on large sets of various features with several parameters for each of them (LBP, Haar-Like, HOG, Region Covariance Descriptor, etc.). The variety of features is motivated by three main reasons:

- Using fast features like LBP and Haar-like in the first levels of the cascade allows a fast rejection of a high part of negatives. The remaining ones will be rejected by a more sophisticated feature like Covariance Descriptor. This will highly decrease the detection time.
- Covariance Descriptor are not discriminative enough for very small regions. Our aim is to train the new detector on specific body parts, especially the upper one (shoulders and heads) to increase detection rate in highly crowded scenes (with a high rate of partial occlusions). Using a large set of various features allows the training system to select the ones which provide the best discriminative power for these regions.
- The possibility to combine several features to describe the same region, even by a simple concatenation, providing more discriminative power than using single features.

Another part of this approach consists in the optimization of the detector at two levels:

- Optimizing the training process by first clustering both positive and negative training samples. This clustering allows to focus on the hard samples which are too close to the other class from a classification point of view, providing more accurate detectors.
- Iterative training of several detectors on randomly selected samples, and weighting of the training samples according to their classification confidence, which allows to improve the clustering process.

The evaluation of this approach is still in progress.

6.4. Walking Speed Detection on a Treadmill using an RGB-D camera : experimentations and results

Participants: Baptiste Fosty, François Brémond.

keywords: RGB-D camera analysis, walking speed, serious games

Within the context of the developement of serious games for people suffering from Alzheimer disease (Az@Game project), we have developed an algorithm to compute the walking speed of a person on a treadmill. The goal is to use this speed inside the game to control the displacement of the avatar, and then for the patient to perform some physical as well as cognitive tasks. For the evaluation of the accuracy of the algorithm, we collected a video data set of healthy people walking on a motorized treadmill.

Figure 10. Some examples of detection using the previously proposed approach (see section Early Work).

Protocol. With the help of a specialist in the domain of physical activities, a protocol has been set up to cover the spectrum of the possible walking speeds and to prove the reproducibility of the results. This protocol consists in performing three times ten minutes of walking on the motorized treadmill, each attempt being itself divided in five times two minutes at the following speeds : 1.5 km/h, 2.5 km/h, 3.5 km/h, 4.5 km/h and 5.5 km/h. Participants, mostly people from the age of 18 to 60 without any physical disorder that could influence the gait, were asked to keep a natural gait and to follow the rotation of the treadmill.

Ground truth. The quantitative performances of the walking speed computation are evaluated by comparison with the speed of the walking person. The speed references are twofold :

- a theoretical value : the speed displayed by the treadmill, set up by the participant but imposed by the protocol (see Figure 12 , red graph),
- a practical value : white marks have been painted on the treadmill to recompute the real speed of the rotation and so the walking speed (see Figure 12 , green and blue graphs).

Results. The results presented herein are based on the videos of 36 participants who performed the protocol described above, with 17 males and 19 females, with an average age of 32.1±7.7 years, an average height of 171.1±9.1 cm and an average weight of 67.4±13.6 kg.

Speed (km/h)	Number of observations	Average (km/h)	Standard deviation (km/h)	Mean absolute error (km/h)
1.5 (WM)	11823 (3170)	1.59 (1.508)	0.20 (0.014)	0.15 (0.011)
2.5 (WM)	14549 (5113)	2.56 (2.478)	0.20 (0.026)	0.16 (0.027)
3.5 (WM)	17399 (7133)	3.53 (3.453)	0.20 (0.040)	0.16 (0.051)
4.5 (WM)	19800 (9129)	4.47 (4.427)	0.21 (0.059)	0.17 (0.079)
5.5 (WM)	22163 (11180)	5.36 (5.397)	0.26 (0.116)	0.21 (0.116)
Total (WM)	**85734 (35725)**	**3.5 (3.453)**	**0.22 (0.074)**	**0.17 (0.071)**

Figure 11. Walking speed results. The number of observations corresponds to the number of time the speed has been computed. The WM lines refer to the results of the detection of the white mark.

The table in figure 11 shows the statistical evaluation of the performances of the system. The average column shows that the accuracy of the system is better for the median speeds (around 4.5km/h). When the person is walking slower, the system overestimate the speed due to the wrongly detected steps whereas when faster, there is an underestimation because of missing the exact time when the distance between feet is maximum (framerate too low).
A paper reporting this work is actually under writing process.

6.5. Head Detection Using RGB-D Camera

Participants: Marine Chabran, François Brémond.

keywords: RGB-D camera analysis, head detection, serious games

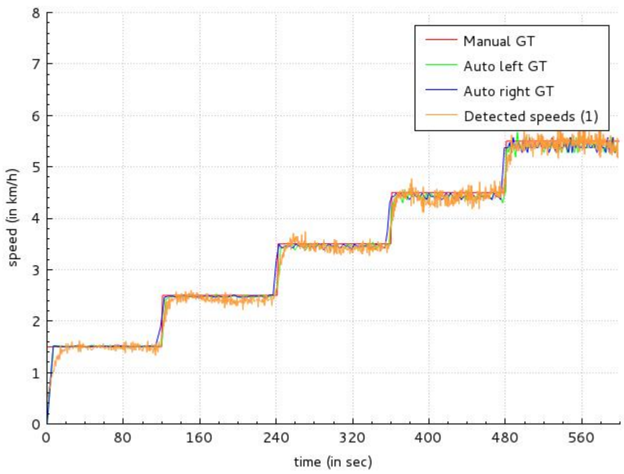


Figure 12. Graphical results of one attempt of protocal issue. The red graph corresponds to the theoretical speed, green and blue graphs correspond respectively to the detected speed of the left and right white marks on the treadmill, and the orange graph is the detected walking speed of the person.

The goal of this work is to improve a head detection algorithm using RGB-D sensor (like a Kinect camera) for action recognition as part of a study of autism. The psychologists want to compare the learning process of children with autism syndrome depending on games (digital or physical toys).

The algorithm described in [79] represents a head by its center position. It takes three steps to determine this point :

- Determine possible head center positions using a head model : inner circle radius=6 cm, outer circle radius=20 cm (Figure 13).
  A good inner point is a point on the inner circle verifying :

  $$depthHeadCenter + 30cm > depthInnerPoint > depthHeadCenter - 30cm.$$

  A good outer point is a point on the outer circle verifying :

  $$depthHeadCenter < depthOuterPoint + 15cm.$$

- Merge close head centers separated by less than 4 pixels.
- Select final head center according to its score (calculated according to the number of good inner and outer points).

For now, it works well within video where people are close to the camera (about 1 meter) and without any background just behind them (Figure 14).

The problem is when the person is sitting and the head is ahead of the body (Figure 15) or close to a wall, the difference between head depth and outer circle depth becomes not sufficient (about 10 cm).

We have evaluated the performance of this algorithm with two data sets (Table 1). For Lenval Hospital data set, we have evaluated 2 series of 200 frames, for the Smart Home data set, we have evaluated 3 series of 300 frames (a total of 1300 heads).

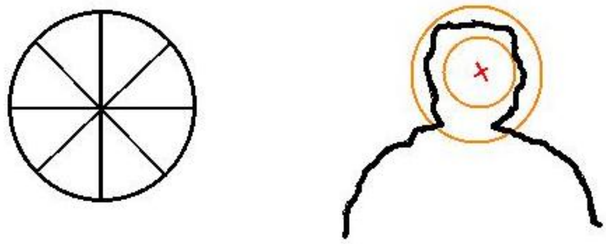


Figure 13. Each circle is divided in n parts (n=8). The points on the inner circle must have a similar depth with the center point, the points on the outer circle must be further than the center point compared to the camera

Table 1. Performance of head detection and people detection on two different data sets.

Videos	Head Detection (%)	People detection (%)
Lenval Hospital dataset (Figure 14)	89.7	96.9
Rest home dataset (Figure 15)	62.8	85.3

Figure 14. Result of head detection - good detection. The bounding box represents the person, the small blue circle represents the head center.

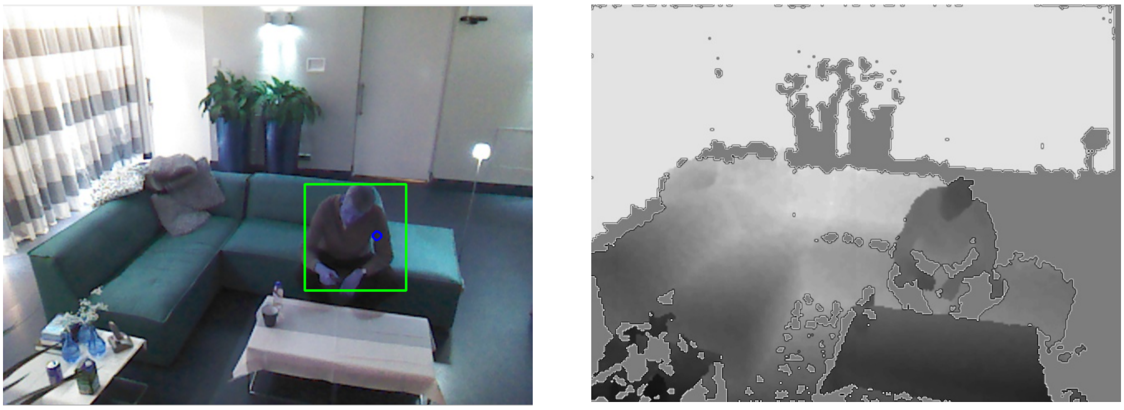


Figure 15. Result of head detection - wrong detection. The green bounding box represents the person, the small blue circle represents the head center.

6.6. Video Segmentation and Multiple Object Tracking

Participants: Ratnesh Kumar, Guillaume Charpiat, Monique Thonnat.

keywords: Fibers, Graph Partitioning, Message Passing, Iterative Conditional Modes, Video Segmentation, Video Inpainting

This year we focussed on multiple object tracking, and writing of the thesis manuscript of Ratnesh (defense on December 2014).

The first contribution of this thesis is in the domain of video segmentation wherein the objective is to obtain a dense and coherent spatio-temporal segmentation. We propose joining both spatial and temporal aspects of a video into a single notion *Fiber*. A Fiber is a set of trajectories which are spatially connected by a mesh. Fibers are built by jointly assessing spatial and temporal aspects of the video. Compared to the state-of-the-art, a fiber based video segmentation presents advantages such as a natural spatio-temporal neighborhood accessor by a mesh, and temporal correspondences for most pixels in the video. Furthermore, this fiber-based segmentation is of quasi-linear complexity w.r.t. the number of pixels. The second contribution is in the realm of multiple object tracking. We proposed a tracking approach which utilizes cues from point tracks, kinematics of moving objects and global appearance of detections. Unification of all these cues is performed on a Conditional Random Field. Subsequently this model is optimized by a combination of message passing and an Iterated Conditional Modes (ICM) variant to infer object-trajectories. A third, minor, contribution relates to the development of suitable feature descriptor for appearance matching of persons. All of our proposed approaches achieve competitive and better results (both qualitatively and quantitatively) than state-of-the-art open source datasets.

This first part of the thesis was published at IEEE WACV at the beginning of this year [43], and the work on *multiple object tracking* was recently presented at Asian Conference on Computer Vision [44]

Sample visual results from our recent publication [44] can be seen in Figure 16 .

6.7. Enforcing Monotonous Shape Growth or Shrinkage in Video Segmentation

Participant: Guillaume Charpiat [contact].

This work has been done in collaboration with Yuliya Tarabalka (Ayin team, Inria-SAM), Bjoern Menze (Technische Universität München, Germany), and Ludovic Brucker (NASA GSFC, USA) [http://www.nasa.gov].

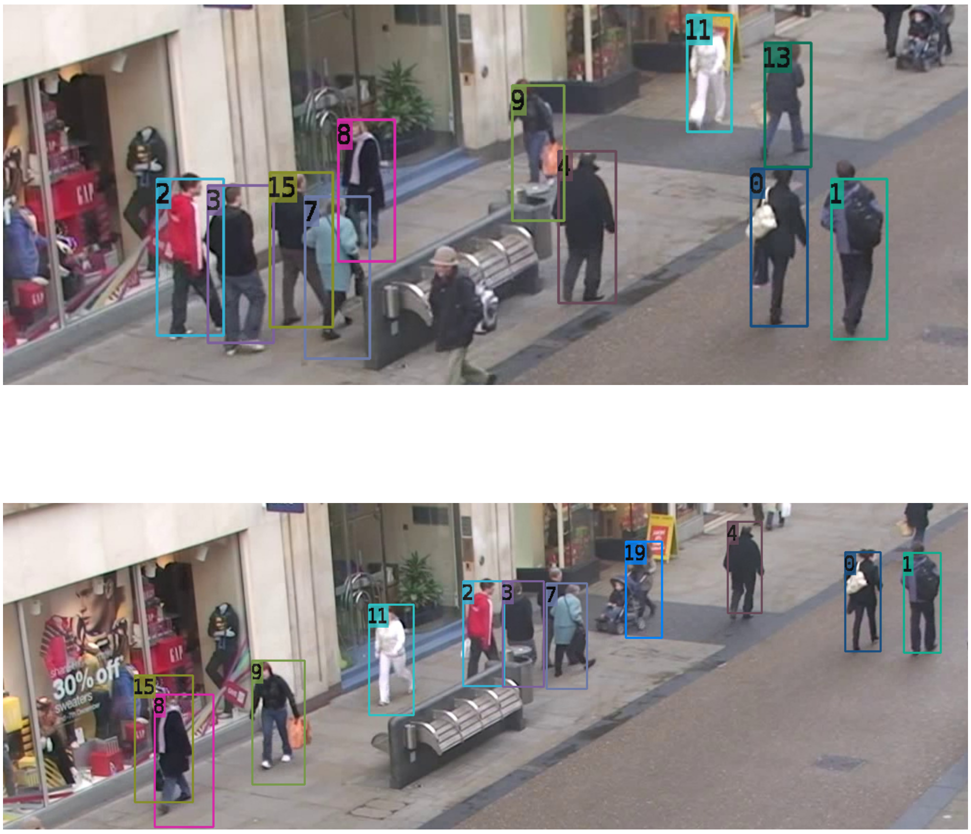


Figure 16. ***Consistent people crossing in dense scenarios.*** *The two images are 121 frames apart.*

keywords: Video segmentation, graph cut, shape analysis, shape growth

The automatic segmentation of objects from video data is a difficult task, especially when image sequences are subject to low signal-to-noise ratio or low contrast between the intensities of neighboring structures. Such challenging data are acquired routinely, for example, in medical imaging or satellite remote sensing. While individual frames can be analyzed independently, temporal coherence in image sequences provides a lot of information not available for a single image. In this work, we focused on segmenting shapes that grow or shrink monotonically in time, from sequences of extremely noisy images.

We proposed a new method for the joint segmentation of monotonically growing or shrinking shapes in a time sequence of images with low signal-to-noise ratio [32]. The task of segmenting the image time series is expressed as an optimization problem using the spatio-temporal graph of pixels, in which we are able to impose the constraint of shape growth or shrinkage by introducing unidirectional infinite-weight links connecting pixels at the same spatial locations in successive image frames. The globally-optimal solution is computed with graph-cuts. The performance of the proposed method was validated on three applications: segmentation of melting sea ice floes; of growing burned areas from time series of 2D satellite images; and of a growing brain tumor from sequences of 3D medical scans. In the latter application, we imposed an additional inter-sequences inclusion constraint by adding directed infinite-weight links between pixels of dependent image structures. Figure 17 shows a multi-year sea ice floe segmentation result. The proposed method proved to be robust to high noise and low contrast, and to cope well with missing data. Moreover, in practice, its complexity was linear in the number of images.

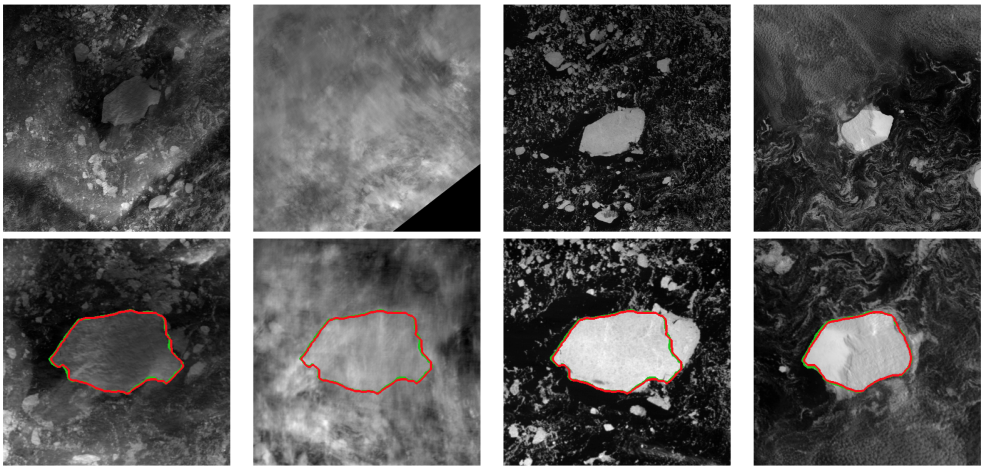


Figure 17. Top: MODIS images at four time moments (days 230, 233, 235 and 267 of 2008, respectively). Bottom: corresponding aligned images with segmentation contours (in red). Manual segmentation (ground truth) is shown in green.

6.8. Multi-label Image Segmentation with Partition Trees and Shape Prior

Participant: Guillaume Charpiat [contact].

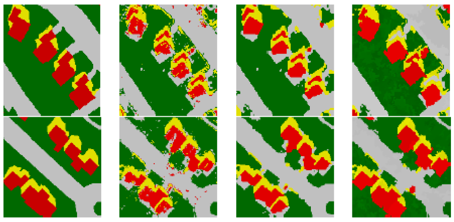
This work has been done in collaboration with Emmanuel Maggiori and Yuliya Tarabalka (Ayin team, Inria-SAM).

keywords: partition trees, multi-class segmentation, shape priors, graph cut

The multi-label segmentation of images is one of the great challenges in computer vision. It consists in the simultaneous partitioning of an image into regions and the assignment of labels to each of the segments. The problem can be posed as the minimization of an energy with respect to a set of variables which can take one of multiple labels. Throughout the years, several efforts have been done in the design of algorithms that minimize such energies.

We propose a new framework for multi-label image segmentation with shape priors using a binary partition tree [50]. In the literature, such trees are used to represent hierarchical partitions of images, and are usually computed in a bottom-up manner based on color similarities, then processed to detect objects with a known shape prior. However, not considering shape priors during the construction phase induces mistakes in the later segmentation. This study proposes a method which uses both color distribution and shape priors to optimize the trees for image segmentation. The method consists in pruning and regrafting tree branches in order to minimize the energy of the best segmentation that can be extracted from the tree. Theoretical guarantees help reducing the search space and make the optimization efficient. Our experiments (see Figure 18) show that the optimization approach succeeds in incorporating shape information into multi-label segmentation, outperforming the state-of-the-art.

(a) Color image. (b) Manual segment. (c) BPT optimization.



(d) Manual. (e) SVM. (f) GC. (g) BPT Opt.

	$\mathcal{A}$	$\mathcal{D}$
SVM	0.764	0.50
GC	0.799	0.49
TC	0.804	0.47
Opt.	0.813	0.69

(h) Performance.

Time [s]	Time [s] (Props. 1-2)
214	84

(i) Execution time.

Figure 18. Classification results for the satellite image over Brest. $\mathcal{A}$ denotes overall classification accuracy, and $\mathcal{D}$ denotes average building's overlap. The performance of the proposed binary partition tree (BPT) optimization method is compared with the following methods: 1) support vector machines (SVM) classification; 2) graph cut (GC) with α-expansion; 3) cut on the BPT, regularized by the number of regions without using shape priors (TC).

6.9. Automatic Tracker Selection and Parameter Tuning for Multi-object Tracking

Participants: Duc Phu Chau, Slawomir Bak, François Brémond, Monique Thonnat.

Keywords: object tracking, machine learning, tracker selection, parameter tuning

Many approaches have been proposed to track mobile objects in a scene [87], [45]. However the quality of tracking algorithms always depends on video content such as the crowded level or lighting condition. The selection of a tracking algorithm for an unknown scene becomes a hard task. Even when the tracker has already been determined, there are still some issues (e.g. the determination of the best parameter values or the online estimation of the tracking reliability) for adapting online this tracker to the video content variation. In order to overcome these limitations, we propose the two following approaches.

The main idea of the first approach is to learn offline how to tune the tracker parameters to cope with the tracking context variations. The tracking context of a video sequence is defined as a set of six features: density of mobile objects, their occlusion level, their contrast with regard to the surrounding background, their contrast variance, their 2D area and their 2D area variance. In an offline phase, training video sequences are classified by clustering their contextual features. Each context cluster is then associated to satisfactory tracking parameters using tracking annotation associated to training videos. In the online control phase, once a context change is detected, the tracking parameters are tuned using the learned parameter values. This work has been published in [30].

A limitation of the first approach is the need of annotated data for training. Therefore we have proposed a second approach without training data. In this approach, the proposed strategy combines an appearance tracker and a KLT tracker for each mobile object to obtain the best tracking performance (see figure 19). This helps to better adapt the tracking process to the spatial distribution of objects. Also, while the appearance-based tracker considers the object appearance, the KLT tracker takes into account the optical flow of pixels and their spatial neighbours. Therefore these two trackers can improve alternately the tracking performance.

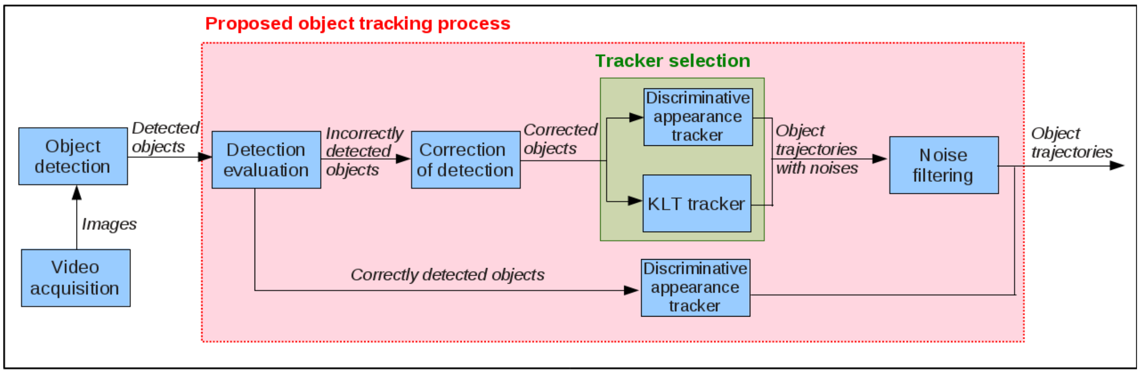


Figure 19. The scheme of the second approach

The second approach has been experimented on three public video datasets. Figure 20 presents correct tracking results of this approach even with strong object occlusion in PETS 2009 dataset. Table 2 presents the evaluation results of the proposed approach, the KLT tracker, the appearance tracker and different trackers from the state of the art. While using separately the KLT tracker or the appearance tracker, the performance is lower than other approaches from the state of the art. The proposed approach by combining these two trackers improves significantly the tracking performance and obtains the best values for both metrics. This work has been published in [39].

Table 2. Tracking results on the PETS sequence S2.L1, camera view 1, sequence time 12.34. The best values are printed in bold.

Method	MOTA	MOTP
Berclaz et al. [60]	0.80	0.58
Shitrit et al. [86]	0.81	0.58
KLT tracker	0.41	**0.76**
Appearance tracker	0.62	0.63
Proposed approach	**0.86**	0.72

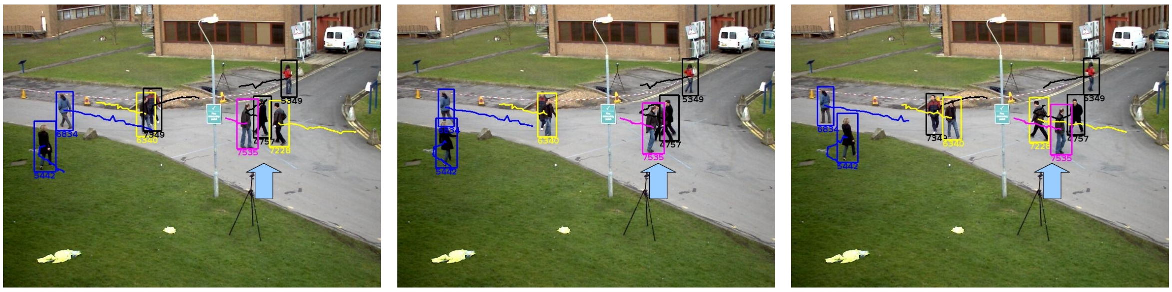


Figure 20. Tracker result: Three persons of Ids 7535, 7228 and 4757 (marked by the cyan arrow) are occluded each other but their identities are kept correctly after occlusion.

6.10. An Approach to Improve Multi-object Tracker Quality Using Discriminative Appearances and Motion Model Descriptor

Participants: Thi Lan Anh Nguyen, Duc Phu Chau, François Brémond.

Keywords: Tracklet fusion, Multi-object tracking

Many recent approaches have been proposed to track multi-objects in a video. However, the quality of trackers is remarkably effected by video content. In the state of the art, several algorithms are proposed to handle this issue. The approaches in [39] and [64] propose methods which compute online or learn descriptor weights during tracking process. These algorithms adapt the tracking to the scene variations but are less effective when mis-detection occurs in a long period of time. Inversely, the algorithms in [59] and [58] can recover a long-term mis-detection by fusing tracklets. However, the descriptor weights in these tracklet fusion algorithms are fixed in the whole video. Furthermore, above algorithms track objects based on object appearance which is not reliable enough when objects look similar to each other.

In order to overcome mentioned issues, the proposed approach brings three contributions: (1) appearance descriptors and motion model combination, (2) online discriminative descriptor weight computation and (3) discriminative descriptors based tracklet fusion. In particular, the appearance of one object can be discriminative with other objects in this scene but can be similar with other objects in another scene. Therefore, tracking objects based on only object appearance is less effective. In order to improve tracker quality, assuming that objects move with constant velocity, this approach firstly combines a constant velocity model from [70] and other appearance descriptors. Continuously, discriminative descriptor weights are computed online to adapt the tracking to each video scene. The more a descriptor discriminates one tracklet over other tracklets, the higher its weight value is. Next, based on these descriptor weights, the similarity score between the target tracklet with its candidate is computed. In the last step, tracklets are fused to a long trajectory by Hungarian algorithm with the optimization of global similarity scores.

The proposed approach gets results of tracker in [63] as input and is tested on challenge datasets. This approach achieves comparable results with other trackers from the state of the art. Figure 1 shows that the tracklet keeps its ID even when occlusion occurs. Table 1 shows the better performance of this approach compared to other trackers from the state of the art.

6.11. Person Re-identification by Pose Priors

Participants: Slawomir Bak, Sofia Zaidenberg, Bernard Boulay, Filipe Martins, Francois Brémond.

keywords: re-identification, pose estimation, metric learning

Human appearance registration, alignment and pose estimation

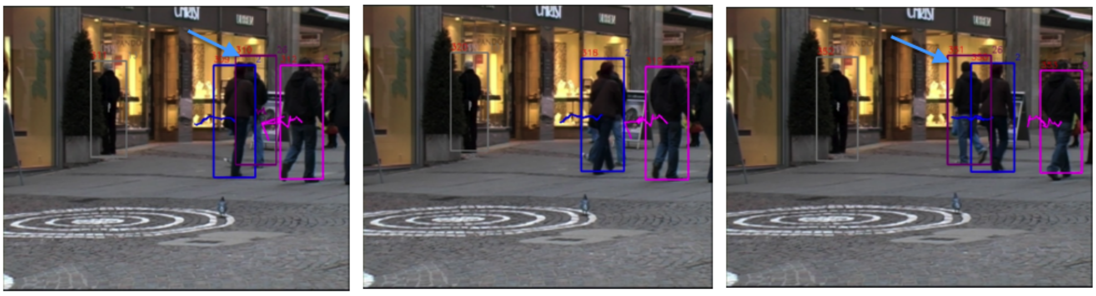
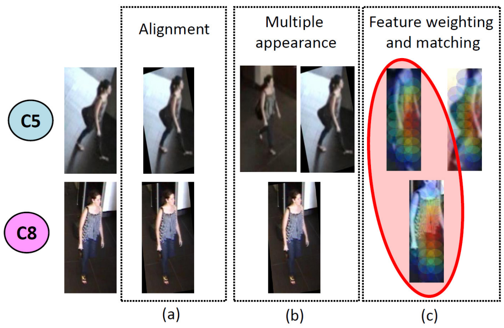


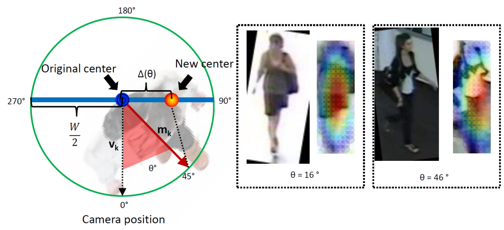
Figure 21. The proposed approach with TUD dataset: Object ID_{26} (presented by pink bounding box) keeps its ID correctly after 11 frames of mis-detection.

Table 3. Tracking results on datasets: TUD-Stadtmitte and TUD-crossing. The best values are printed in bold

Dataset	Method	MT(%)	PT(%)	ML(%)
TUD-Stadtmitte	[57]	60.0	30.0	10.0
TUD-Stadtmitte	[30]	**70.0**	10.0	20.0
TUD-Stadtmitte	[71]	**70.0**	30.0	**0.0**
TUD-Stadtmitte	[95]	**70.0**	30.0	**0.0**
TUD-Stadtmitte	Ours	**70.0**	30.0	**0.0**
TUD-Crossing	[89]	**53.8**	38.4	7.8
TUD-Crossing	Ours	**53.8**	46.2	**0.0**



(a) left



(b) right

Figure 22. Improvements on re-identification using viewpoint cues: (a) target alignment; (b) multiple target appearance based on clustering; (c) pose orientation-driven weighting. The left illustration shows an example of the same person viewed from two different cameras. The right image presents pose estimation algorithm.

Re-identifying people in a network of cameras requires an invariant human representation. State of the art algorithms are likely to fail in real-world scenarios due to serious perspective changes. Most of existing approaches focus on invariant and discriminative features, while ignoring the body alignment issue. In this work we proposed 3 methods for improving the performance of person re-identification. We focus on eliminating perspective distortions by using 3D scene information. Perspective changes are minimized by affine transformations of cropped images containing the target (1). Further we estimate the human pose for (2) clustering data from a video stream and (3) weighting image features. The pose is estimated using 3D scene information and motion of the target. Pose orientation is computed by dot product between viewpoint vector and motion of the target (see figure 22). We validated our approach on a publicly available dataset with a network of 8 cameras. The results demonstrated significant increase in the re-identification performance over the state of the art [36].

Matching employing pose priors

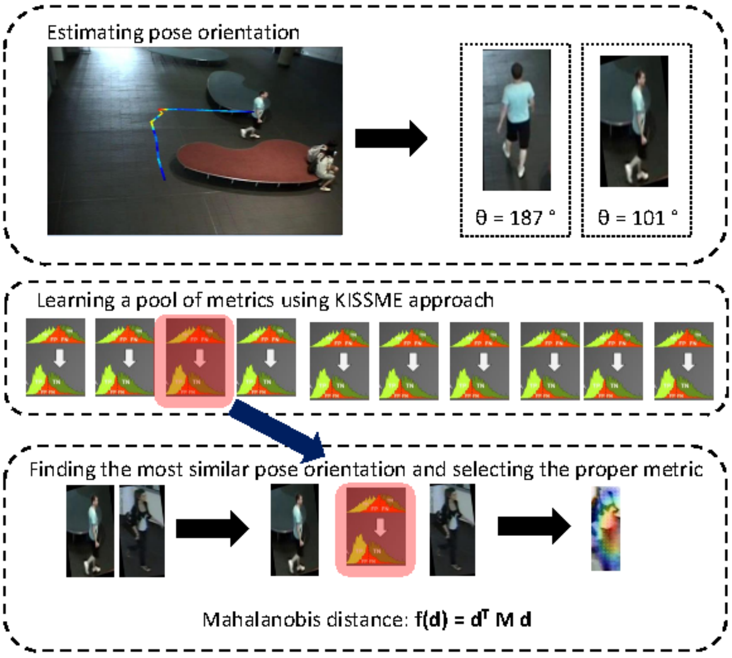


Figure 23. The proper metric is selected from the pool of previously learned metric for matching different poses.

Currently we are working on learning the matching strategy of appearance extracted from different poses. We employ well known metric learning tools for matching given poses. Let us assume that pose can be described by the angle between the motion vector of the target and the viewpoint vector of the camera (see figure 22). Thus for each target appearance we can express the pose as the angle in the range of [0,360). We decide to divide this range into n bins. Given n bins of estimated poses, we learn how to match different poses corresponding to different bins. In the result, we learn $n * (n + 1)/2$ metrics. While learning metrics, we follow a well known scheme based on image pairs, containing two different poses of the same target as positives and pairs of different poses containing different targets as negatives. The learned metrics stand for the metric pool. This metric pool is learned offline and does not depend on camera pair. In the result, once metric pool is learned, it can be used for any camera pair.

Given two images from different (or the same) camera, we first estimate the poses for each image. Having two poses, we select a corresponding metric from the metric pool. The selected metric provides the strategy to compute similarity between two images (see figure 23).

6.12. Global Tracker : An Online Evaluation Framework to Improve Tracking Quality

Participants: Julien Badie, Slawomir Bak, Duc Phu Chau, François Brémond, Monique Thonnat.

keywords: online quality estimation, re-identification, tracking results improvements

This work addresses the problem of estimating the reliability of a tracking algorithm during runtime and correcting the anomalies found. Evaluating and tuning a tracking algorithm generally requires multiple runs and ground truth. The proposed framework called global tracker overcomes these limitations by combining an online evaluation algorithm and a recovering post-process.

Designing an evaluation framework that does not require ground truth has many different applications. One of them is to provide feedback to the tracking algorithm that can tune its own parameters to improve the results on the next frame. Another convenient application is to filter the reliable information from the tracking algorithm that can be used by the next processing step such as event recognition or re-identification.

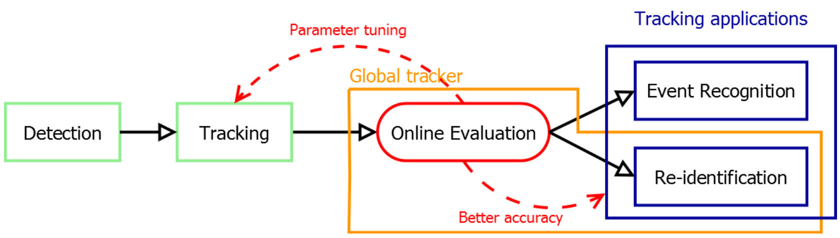


Figure 24. Integration of the global tracker, combining online evaluation and re-identification

The proposed online evaluation framework is based on control features. It means that several representative values or characteristics (the features) are chosen to monitor what is happening. Control features are the features that the online evaluation framework uses to monitor the status of the tracked objects. The framework is divided into two steps :

- computing the control features related to each tracked object of the scene on the current frame
- detecting the possible anomalies and then classifing them into two categories : anomalies due to tracking errors or benign anomalies (when a person leaves the scene or cross an obstacle for example).

This approach has been tested on two datasets (PETS 2009 and Caviar) with two different tracking algorithms (a multi-feature tracker and a tracker based on graph partitioning). The results show that the global tracker, even associated with a tracking algorithm that does not have good results, can perform as well as the state of the art.

This approach has been published in AVSS 2014 [33] which details the differences between real errors and benign anomalies.

6.13. Human Action Recognition in Videos

Participants: Piotr Bilinski, François Brémond.

keywords: Action Recognition; Human Action Recognition

Table 4. Tracking results on sequence S2.L1.View1 of the PETS2009 dataset using CLEAR metrics

Methods	MOTA	MOTP
Heili et al. [72]	0.89	0.66
Zamir et al. [96]	**0.90**	0.69
Milan et al. [77]	**0.90**	**0.74**
Tracker 1	0.62	0.63
Tracker 1 + global tracker	0.85	0.71
Tracker 2	0.85	**0.74**
Tracker 2 + global tracker	**0.90**	**0.74**

Table 5. Tracking results on the Caviar dataset using Mostly Tracked (MT), Partially Tracked (PT) and Mostly Lost (ML) metrics

Method	MT (%)	PT (%)	ML (%)
Li et al. [76]	84.6	14.0	1.4
Kuo et al. [74]	84.6	14.7	**0.7**
Tracker 1	78.3	16.0	5.7
Tracker 1 + global tracker	**86.4**	**8.3**	5.3

This Ph.D. thesis targets the automatic recognition of human actions in videos. Human action recognition is defined as a requirement to determine what human actions occur in videos. This problem is particularly hard due to enormous variations in visual and motion appearance of people and actions, camera viewpoint changes, moving background, occlusions, noise, and enormous amount of video data.

Firstly, we review, evaluate, and compare the most popular and the most prominent state-of-the-art techniques, and we propose our action recognition framework based on local features, which we use throughout this thesis work embedding the novel algorithms. Moreover, we introduce a new dataset (CHU Nice Hospital) with daily self care actions of elder patients in a hospital.

Then, we propose two local spatio-temporal descriptors for action recognition in videos. The first descriptor is based on a covariance matrix representation, and it models linear relations between low-level features. The second descriptor is based on a Brownian covariance, and it models all kinds of possible relations between low-level features.

Then, we propose three higher-level feature representations to go beyond the limitations of the local feature encoding techniques.

The first representation is based on the idea of relative dense trajectories. We propose an object-centric local feature representation of motion trajectories, which allows to use the spatial information by a local feature encoding technique.

The second representation encodes relations among local features as pairwise features. The main idea is to capture the appearance relations among features (both visual and motion), and use geometric information to describe how these appearance relations are mutually arranged in the spatio-temporal space.

The third representation captures statistics of pairwise co-occurring visual words within multi-scale feature-centric neighbourhoods. The proposed contextual features based representation encodes information about local density of features, local pairwise relations among the features, and spatio-temporal order among features.

Finally, we show that the proposed techniques obtain better or similar performance in comparison to the state-of-the-art on various, real, and challenging human action recognition datasets (Weizmann, KTH, URADL, MSR Daily Activity 3D, HMDB51, and CHU Nice Hospital).
The Ph.D. thesis was defended on December 5, 2014.

6.14. Action Recognition Using 3D Trajectories with Hierarchical Classifier

Participants: Michal Koperski, Piotr Bilinski, François Brémond.

keywords: action recognition, computer vision, machine learning, 3D sensors

The goal of our work is to extend recently published approaches ([61], [93]) for Human Action Recognition to take advantage of the depth information from 3D sensors.

We propose to add depth information to trajectory based algorithms ([61], [93]). Currently mentioned algorithms compute trajectories by sampling video frames and then tracking points of interest - creating the trajectory. Our contribution is to create even more discriminative features by adding depth information to previously detected trajectories. In our work we propose methods to deal with noise and missing measurements in depth map.

The second contribution is a technique to deal with actions which do not contain enough motion to compute discriminative trajectory descriptors. Actions like sitting, standing, laptop use do not contain large amount of motion, or motion is occluded by the object. For such cases we proposed LDP (Local Depth Pattern) descriptor which does not require motion to be computed.

Proposed descriptors are further processed using a Bag of Words method and SVM classifier. We use hierarchical approach where at first level we train classifier to recognize if given example contains high or low amount of motion. Then at second layer we train SVM classifier to recognize action labels.

Figure 25. Visualization of MSR Dailiy Activty 3D data set (left) - video input frame, (center) - frame with detected trajectories (red - static points, green detected trajectories, (right) - corresponding depth map

The evaluation of our method was conducted on "Microsoft Daily Activity3D" data set [94] which consists of 16 actions (drink, eat, read book, call cellphone, write on a paper, use laptop etc.) performed by 10 subjects. We achieve superior performance among techniques which do not require skeleton detection. This work was published in proceedings of the 21st IEEE International Conference on Image Processing, ICIP 2014 [42]

6.15. Action Recognition using Video Brownian Covariance Descriptor for Human

Participants: Piotr Bilinski, Michal Koperski, Slawomir Bak, François Brémond.

keywords: action recognition, computer vision, machine learning

This work addresses a problem of recognizing human actions in video sequences. Recent studies have shown that methods which use bag-of-features and space-time features achieve high recognition accuracy [61], [93], [42]. Such methods extract both appearance-based and motion-based features. In image processing, a novel trend has emerged that ignores explicit values of given features, focusing instead on their pairwise relations. The most known example of such an approach is covariance descriptor [92]. Inspired by Brownian motion statistics [88] and application in people Re-identification [35]; we propose to model relationships between different pixel-level appearance features such as intensity and gradient using Brownian covariance, which is a natural extension of classical covariance measure. While classical covariance can model only linear relationships, Brownian covariance models all kinds of possible relationships. We propose a method to compute Brownian covariance on space-time volume of a video sequence. We show that proposed Video Brownian Covariance (VBC) descriptor carries complementary information to the Histogram of Oriented Gradients (HOG) descriptor. The fusion of these two descriptors gives a significant improvement in performance on three challenging action recognition datasets. The result of this work was published in proceedings of the 11th IEEE International Conference on Advanced Video and Signal-Based Surveillance, AVSS 2014 [38].

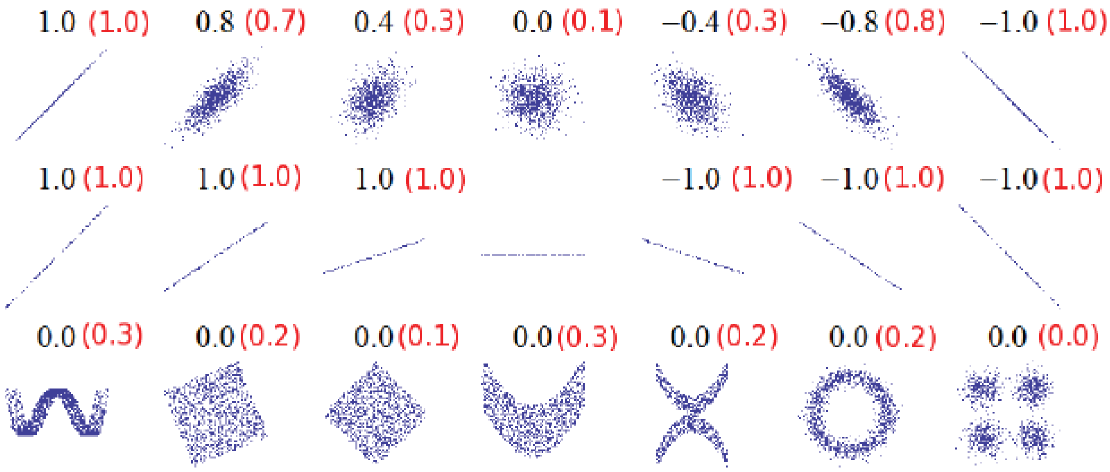


Figure 26. Comparison between Covariance and Brownian distance correlation. Covariance values in black, Brownian values in red.

6.16. Towards Unsupervised Sudden Group Movement Discovery for Video Surveillance

Participants: Sofia Zaidenberg, Piotr Bilinski, François Brémond.

keywords: Event detection; Motion estimation; Anomaly estimation; Situation awareness; Scene Understanding; Group Activity Recognition; Stream Selection

We present a novel and unsupervised approach for discovering "sudden" movements in surveillance videos. The proposed approach automatically detects quick motions in a video, corresponding to any action. A set of possible actions is not required and the proposed method successfully detects potentially alarm-raising actions without training or camera calibration. Moreover, the system uses a group detection and event recognition framework to relate detected sudden movements and groups of people, and to provide a semantical interpretation of the scene. We have tested our approach on a dataset of nearly 8 hours of videos recorded from

two cameras in the Parisian subway for a European Project. For evaluation, we annotated 1 hour of sequences containing 50 sudden movements.
This work has been published in [47].

6.17. Autonomous Monitoring for Securing European Ports

Participants: Vasanth Bathrinarayanan, François Brémond.

Keywords: Event Recognition, Port Surveillance

This work is done for the European research project SUPPORT (Security UPgrade for PORTs). This project addresses potential threats on passenger life and the potential for crippling economic damage arising from intentional unlawful attacks on port facilities, by engaging representative stakeholders to guide the development of next generation solutions for upgraded preventive and remedial security capabilities in European ports. The overall benefit is securing and efficient operation of European ports enabling uninterrupted flows of cargo and passengers while suppressing attacks on high value port facilities, illegal immigration and trafficking of drugs, weapons and illicit substances.

Scene understanding platform was tested on this new dataset, which has archived footage from past incidents and some acted scenarios. The processing pipeline of algorithms contains camera calibration, background subtraction using GMM (Gaussian Mixture Model), people detection using DPM (Deformable Parts Model), Tracking (Frame to Frame), Event recognition.

We collected several hours of videos which contained security related events like Intrusion to port by different methods (sea, gates, fences), Spying activities from outside the port, robbery or theft, ticketless travelling, restricted zone access, abondon luggage and some abnormal behaviors. The system was modelled and validated for all the above events to be detected and also a live real time demo was done for the completion of the project. All the events from our systems are later sent to project partners for fusion of the data with other sensors data and information from police, internal and external reports, etc., to detect complex security threats (see figure 27).

6.18. Video Understanding for Group Behavior Analysis

Participants: Carolina Garate, François Brémond.

keywords: Computer vision, group tracking, scene understanding, group behavior recognition, video surveillance, event detection.

The main work in this PhD thesis concerns the recognition of the behaviors of a group of people (2-5 persons) involved in a scene depicted by a video sequence.

Our goal focuses on the automatic recognition of behavior patterns in video sequence for groups of people (2-5 persons). We want to build a real time system able to recognize various group scenarios.

The approach includes different tasks to achieve the final recognition. The first one consists in tracking groups of moving regions detected in the video sequence acquired by the cameras. The second task attempts to classify these moving regions into people classes. Finally, the last task recognizes group scenarios using a priori knowledge containing scenario models predefined by experts and also 3D geometric and semantic information of the observed environment.

Our approach considers a chain process consisting of 5 consecutive steps for video processing. The steps are : 1) segmentation, 2) physical object detection, 3) physical objects tracking, 4) group tracking and 5) group behavior recognition. Our research focuses on the last two phases.

First, group scenarios have been defined (and then recognized) using the general scenario description language. Second, the likelihood of the group scenario recognition has been quantified. Third, machine learning techniques have been investigated to learn and recognize these scenarios.

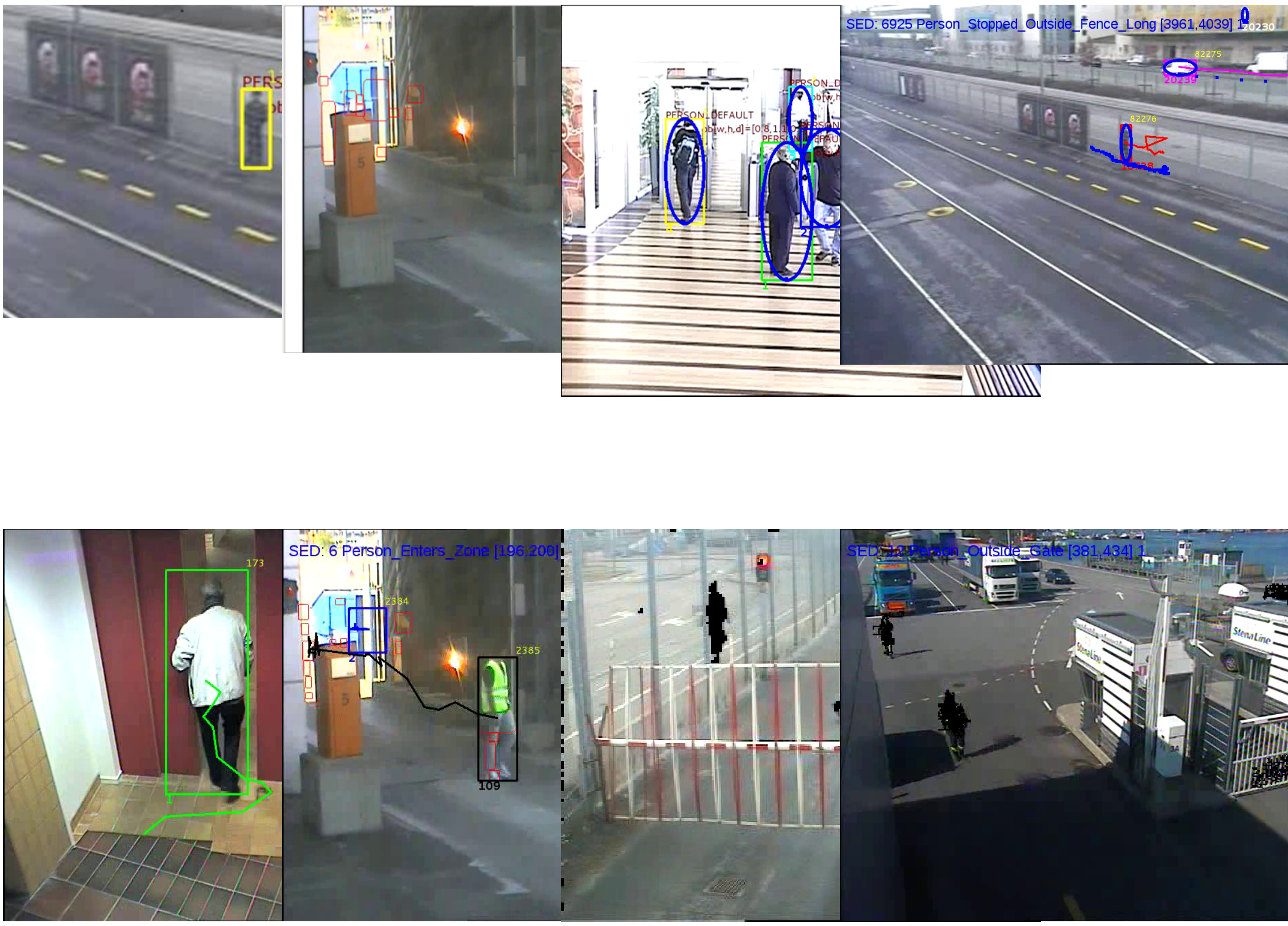


Figure 27. Detection of complex security threats from the tested dataset for Port surveillance

We have processed the data set from 1 month video surveillance camera in the Torino subway and the Minds eye data set. Recognizing several and different events such as: walking groups, standing still groups, running groups, calm groups (i.e. having a bounding box with stable size), active groups (i.e. with bounding box's size variations, meaning that group members move a lot).

6.19. Evaluation of an Event Detection Framework for Older People Monitoring: from Minute to Hour-scale Monitoring and Patients Autonomy and Dementia Assessment

Participants: Carlos F. Crispim-Junior, Alvaro Gomez Uria Covella, Carola Strumia, Baptiste Fosty, Duc Phu Chau, Anh-Tuan Nghiem, Alexandra Konig, Auriane Gros, Philippe Robert, François Brémond.

keywords: RGBD cameras, description-based activity recognition, older people,

Two main works are reported here: the continuous evaluation and extention of our event detection framework for older people monitoring, and the proposal of a behavioral classification model for the assessment of autonomy and cognitive health level of older people using automatically detected events. The evaluation of our event monitoring framework was extended from 29 to 49 recordings of senior participants undertaking physical tasks (7 min per participant, total : 5.71 hours) and instrumental activities of daily living (IADL, 15 minutes per participant, total: 12.25 hours). The recordings have taken place in a ecological observation room set in the Memory Center of Nice hospital. In the extended evaluation we employed a RGBD sensor as input instead of a standard RGB camera due to its advantages like invariance to illumination changes and real-time measurements of 3D information which foster better performance of the underlying algorithms for people detection and tracking. Table 6 presents the event monitoring performance of the present system for 49 participants. Event detection performance on physical task generalized to the larger dataset with a small performance increase of 1.4% (average F-Score). Concerning IADL detection although the global performance value (F-Score, 80.7 %) are the same, the new approach have made a trade-off between recall and precision to obtain more reliable detection of activities and their parameter estimations. Low precision values on preparing drink (e.g., making coffee) and watering plant are due to these activities being performed in very close - if not overlapping - locations (contextual zones). Low precision values in reading are due to the preferred reading location be close to image edges where most parts of person body are frequently outside the camera field of view.

Using the event monitoring system as input we have devised a behavioral classification model for the automatic assessment of participant cognitive health and autonomy level. Besides to event data the model also uses fine-grained data about person gait attributes (*e.g.*, stride-length, cadence, etc), obtained by a RGBD-based algorithm for gait analysis also developed in STARS team. Briefly, the event monitoring system supports the doctor by automatically annotating the patient daily living activities and assessing his/her gait parameters in a quantitative way, and the behavioral model performs the classification of participant's dementia and autonomy levels as a complement for standard psychometric scales for autonomy. We achieved an average accuracy of 83.67 % at the prediction of patient autonomy (poor, mediocre, good), and of 73.46 % for cognitive level class (healhy, memory cognitive impairment - MCI, alzheimer's disease), all models using a Naíve Bayes classifier. The results suggest that the behavioral classification model using automatically detected events outperforms the same model using events manually annotated by domain experts (81 %). On the contrary, the model using annotated data still outperforms the automated detection at dementia classification (79.46 %). Results indicate it is easier to predict the autonomy level than the Dementia, since the latter may be seen as the cause /source and the first its consequences. Deciding whether a decay on cognitive abilities relates to normal aging or early MCI or a given mild cognitive decay is an early symptom of Alzheimer's diasease or a severe case of MCI is also a open-problem for medical community. Future work will focus on investigating whether the remaining performance to achieve is related to the performance failures of the underling event monitoring system, to important behavioral aspects still not covered by the behavioral model, or even to the inherently ambiguous nature of the dementia classes.

Table 6. Event Monitoring Performance

Physical Tasks	Recall	Precision
Single Task	100%	88%
Dual Task	100%	98%
IADLs	Recall	Precision
Preparing drug box	87%	93%
Watering plant	80%	63%
Reading	60%	88%
Prepare drink	90%	68%
Talk on phone	89%	89%

We have also started the evaluation of the event monitoring system in Nursing home scenario passing from a minute time-scale to hours. A first participant was monitored with two RGBD sensors, one for bed-related events (sleep, bed exits) and one for living room and daily living activity events for 14 days. Preliminary results are 80 % for entering in bed and 100 % for bed exit in set of 6 events of each class in 13 hours monitoring (6 pm - 7am). Figure 28 illustrates the detection of restroom usage during the night. The automatic monitoring of participant activities during night is an important contribution to medical/nursing staff as wandering behavior at night is a common cause of accident in older people population. For instance, detecting whether a bed-exit during the night will be followed by a restroom visit or a bedroom exit plays a significant role at predicting a possibly dangerous situation.

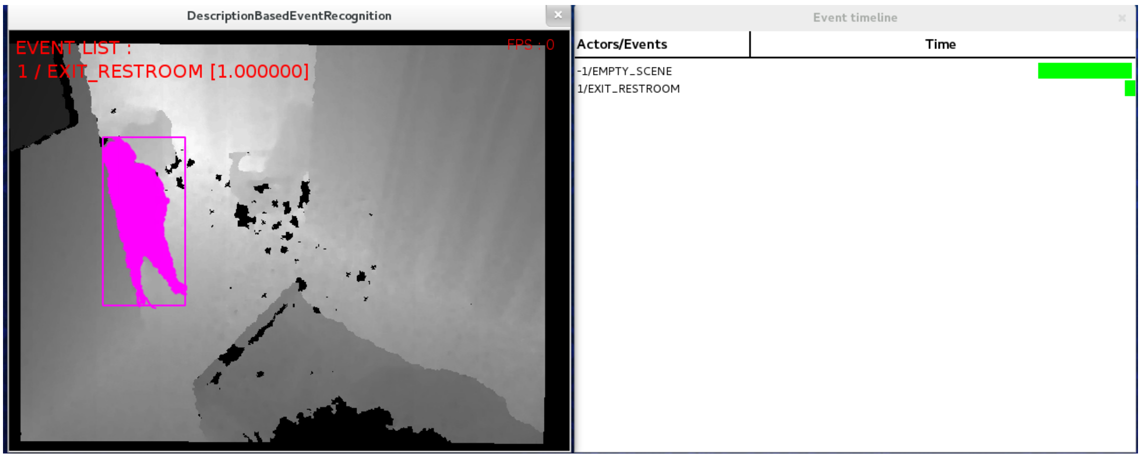


Figure 28. Long-term monitoring of Older People: Sleep-related events

Two papers are envisaged to report the results of this year research to scientific community, one describing the new version of the event monitoring system, and a second one for the developed behavioral classification model. As a publication of this year we hightlight the paper in partnership with Alexandra Konig and Philippe Robert - entitled Validation of an automatic video monitoring system for the detection of instrumental activities of daily living in dementia patients - in the Journal of Alzheimer disease where we summarize the results of the validation of our event monitoring system for the recognition of activities of daily living of participants of Alzheimer's disease study.

6.20. Uncertainty Modeling Framework for Constraint-based Event Detection in Vision Systems

Participants: Carlos F. Crispim-Junior, François Brémond.

keywords: description-based activity recognition, uncertainty modeling, vision system, older people

Event detection has advanced significantly in the past decades relying on pixel- and feature-level representations of video-clips. Although effective, those representations have difficulty on incorporating scene semantics. Alternatively, ontology and description-based approaches for event modeling can explicitly embed scene semantics, but the deterministic nature of such languages is susceptible to noise from underlying components of vision systems. We have developed a probabilistic framework to handle uncertainty on our constraint-based ontology framework for event detection. This task spans from elementary scenarios uncertainty handling (from low-level data and event intra-class variance) to complex scenario semantic modeling, where time ordering in between event sub-components and the effect of missing components (for instance, due to miss-detection) plays a significant role.

Preliminary results of this work have been published in [40], where the presented formalism for elementary event (scenario) uncertainty handling is evaluated on the detection of activities of daily living of participants of the Alzheimer's disease study of Nice hospital using the newest version of our vision system using a RGB-D sensor (Kinect®, Microsoft©) as input. Two evaluations have been carried out: the first one, (a 3-fold cross-validation) focuses on elementary scenario constraint modeling and recognition, and the second one was devoted to complex scenario recognition following a semi-probabilistic approach (n:45).

Table 7 presents the performance of the uncertainty modeling framework on elementary scenario (primitive state) detection for $N:\ 10\ participants; 15\ min.\ each;\ Total:\ 150\ min.$ The 3-fold cross-validation scheme (n:10 participants) is employed for constraint probabilistic distribution learning and event detection evaluation on 10 RGB-D recordings of participants of the Nice hospital clinical protocol for Alzheimer's disease study. "Crisp" term stands for our deterministic constraint-based ontology language for event modeling. Results are reported as the average performance on the crisp and uncertainty frameworks on the validation sets. Results confirm that the uncertainty modeling improves the detection of elementary scenarios in recall (*e.g.,* In zone phone: 84 to 100 %) and precision indices (*e.g.,* In zone Reading: 54.5 to 85.7%).

Table 7. Framework Performance on Elementary Scenario Detection on a 3-fold-cross-validation scheme

	Crisp		Uncertainty	
IADL	Rec.	Prec.	Rec.	Prec.
In zone Pharmacy	100.0	71.4	100	**83.3**
In zone Phone	84.0	95.45	**100.0**	**100.0**
In zone Plant	100.0	81.8	100.0	81.8
In zone Tea	93.3	**77.7**	93.3	73.7
In zone Read	75.0	54.5	75.0	**85.7**

Table 8 presents the performance of the proposed framework on Composite Event Detection for $N:\ 45\ participants;\ 15\ min.\ each;\ Total:\ 675 min.$ Here a hybrid strategy is adopted where the uncertainty modeling is used for elementary scenarios and the crisp constraint-based framework is used for composite event modeling. Results show improvement on recall index of event detection performance, but the uncertainty framework performance on precision index is still worse than the crisp approach. The latter performance may be attributed to the crisp constraints that did not have their uncertainty addressed yet.

Future work will focus on modeling complex scenario constraints such as time ordering and missing components, and on extending the set of low-level uncertainties which are addressed. Moreover, we have been conducting a joint work with partners of Dem@care project to evaluate the uncertainty framework for multiple sensor fusion at decision level. Currently, processed data from different visual modalities (standard RGB, RGBD, and wearable cameras) have been gathered for 17 participants of Nice hospital pilot@lab, and preliminary results are expected for the first semester of 2015.

Table 8. Framework Performance on Composite Event Detection Level

	Crisp		Uncertainty	
IADL	Rec.	Prec.	Rec.	Prec.
Talk on Phone	88.76	89.77	**88.76**	85.86
Preparing Tea/Coffee	81.42	40.36	**92.85**	**55.08**
Using Pharmacy Basket	87.75	95.65	**89.79**	**97.77**
Watering plant	78.57	**84.61**	**100.0**	28.86

6.21. Assisted Serious Game for Older People

Participants: Minh Khue Phan Tran, François Brémond, Philippe Robert.

keywords: interactive system, elderly people, serious game

A system able to interact with older people has been recently devised. The system consists of two parts: Recognition and Interaction. Recognition part, requiring an Asus Xtion Pro Live Camera, consists in observing the scene to decide when is the best moment to interact with users. Afterwards, the Interactive system tries to engage the patient via an interface and through Microsoft Kinect Camera, the patient can interact with the interface using voice or gesture. The interface is designed with Unity 3D game engine (see figure 29).

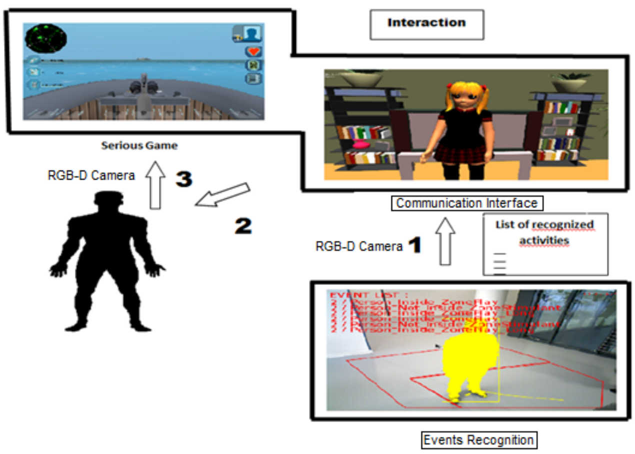


Figure 29. Functional diagram of our interactive system

An experiment was conducted in a memory center for older people, Institut Claude Pompidou in Nice, in order to test different functionalities of the system. Here, participants can experiment the system in a private room (see figure 30) equipped with a large screen and can start the game without having to use devices (mouse, keyboard). The "best moments" to interact with participants are defined when they stay more than 5 seconds in front of the screen. Once these moments are recognized, the interface of Interactive part is called. The avatar indicates the place to be for playing and starts the game.

19 older people have participated to the experiment. 16 succeeded to follow the indications of avatar up to the start of the game. Most of them have appreciated the interaction with the avatar. Even 13 of them have continued to play the second game after suggestion of the avatar.

Future work aims at looking at other indicators (behavior, gaze) that the system can rely on to improve user's interaction.

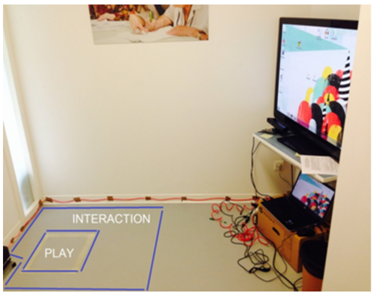


Figure 30. Experimental private room

6.22. Enhancing Pre-defined Event Models Using Unsupervised Learning

Participants: Serhan Coşar, François Brémond.

keywords: Pre-defined activity models, unsupervised learning, tailoring activity models

In this work, we have developed a new approach to recognize human activities from videos, given models that are learned in an unsupervised way and that can take advantage of *a priori* knowledge provided by an expert of the application domain. The description-based methods use pre-defined models and rules to recognize concrete events. But, if the data has unstructured nature, such as daily activities of people, the models cannot handle the variability in data (e.g., the way of preparing meal is person dependent).

In order to overcome this drawback, we have combined the description-based method in [66] with an unsupervised activity learning framework, as presented in Figure 31 . We have created a mutual knowledge loop system, in which both frameworks are combined in a way to compensate their individual limitations. In [66], scene regions are pre-defined and the activity models are created via defining an expected duration value (e.g., 2 seconds) and a posture type (e.g., standing) by hand. Thus, these hand-crafted models fail to cover the variability in data and require an update by experts whenever the scene or person changes. To automatically define these parameters, we utilize the unsupervised activity recognition framework. The unsupervised approach first learns scene regions (zones) in the scene using trajectory information and then, it learns the duration and posture distribution for each zone. By matching the pre-defined zones with learned zones, we connect the learned parameter distributions with hand-crafted models.

The knowledge is passed in a loopy way from one framework to another one. By knowledge we mean: (i) the geometric information and scene semantics of the description-based system are used to label the zones that are learned in an unsupervised way, (ii) the activity models that are learned in an unsupervised way are used to tune the parameters (i.e. tailoring) in the activity models of the description-based framework. It is assumed that the person detection and tracking are already performed and we have the trajectory information of people in the scene beforehand.

We have tested the performance of the knowledge-loop based framework on two datasets: i) Hospital-RGB, ii) Hospital-RGBD. Each dataset contains one person performing everyday activities in a hospital room. The activities considered in the datasets are ”watching TV“, ”preparing tea“, ”answering phone“, ”reading newspaper/magazine“, ”watering plant“, ”organizing the prescribed drugs“, ”writing a check at the office desk“ and ”checking bus routes in a bus map“. Each person is recorded using RGB and RGBD cameras of 640×480 pixels of resolution. RGB dataset consists of 41 videos and RGBD dataset contains 27 videos. For each person, video lasts approximately 15 minutes.

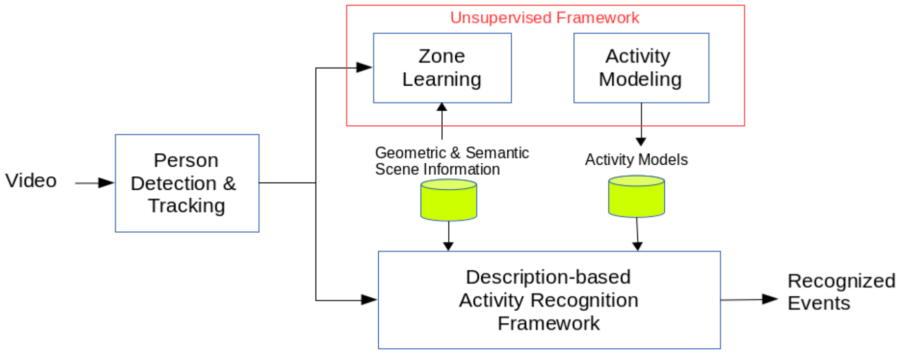


Figure 31. The flow diagram of the framework that combines pre-defined activity models with unsupervised models.

The performance of the approach in [66] with hand-crafted models and our approach with learned models for Hospital-RGB and Hospital-RGBD datasets are presented in Table 9 and in Table 10 . The results have been partially presented in Ellomiietcv2014 (**waiting hal acceptation**). It can be clearly seen that updating the constraints in activity models using data learned by the unsupervised approach enables us to detect activities missed by the pre-defined models. For "watching TV" and "using pharmacy basket" activities in RGB dataset and "answering phone" and "preparing tea" activities in RGBD dataset, there is increase in false positive rates. The reason is that, for some activities, the duration and posture distributions learned by the unsupervised approach can be inaccurate because of other actions occurring inside a zone (e.g., a person standing inside tea zone and reading). For this reason, the constraints updated in activity models are too wide and other activities that occur inside the zone are also detected. Despite the small increase of false positives in some activities, we have increased the true positive rates and obtained sensitivity rates around 90% and 87% on average in RGB and RGBD datasets, respectively, and precision rates around 81% on average in RGBD dataset. Thanks to the distributions learned for time duration and posture, we can enhance the activity models that are manually defined in the description-based, and thereby detect missed events.

Table 9. The activity recognition results of the description-based approach and our knowledge-loop based approach for the Hospital-RGB. The bold values represent the best result for each activity class.

	Hand-crafted Models		Unsupervised Models	
ADLs	Sensitivity (%)	Precision (%)	Sensitivity (%)	Precision (%)
Answering Phone	70	82.35	**95**	**90.47**
Watching TV	84.61	78.57	**100**	54.16
Using Office Desk	91.67	47.82	91.67	**52.38**
Preparing Tea	**80.95**	70.83	76.19	**80**
Using Phar. Basket	100	**90.90**	100	76.92
Watering Plant	100	81.81	88.89	**88.89**
Reading	45.46	83.34	**81.82**	**90**
TOTAL	81.81	**76.52**	**90.5**	76.11

6.23. Using Dense Trajectories to Enhance Unsupervised Action Discovery

Participants: Farhood Negin, Serhan Coşar, François Brémond.

keywords: zone learning, action descriptors, dense trajectories, supervised action recognition, unsupervised activity recognition

Table 10. The activity recognition results of the description-based approach and our knowledge-loop based approach for the Hospital-RGBD datasets. The bold values represent the best result for each activity class.

	Hand-crafted Models		Unsupervised Models	
ADLs	Sensitivity (%)	Precision (%)	Sensitivity (%)	Precision (%)
Answering Phone	80	**100**	**84.21**	88.89
Watching TV	55.56	45.46	**77.78**	**58.34**
Preparing Tea	100	**73.68**	92.85	65
Using Phar. Basket	100	90	100	**100**
Watering Plant	40	66.67	**83.34**	**71.42**
Reading	**100**	66.67	71.42	**83.34**
Using Bus Map	50	71.42	**100**	**100**
TOTAL	75.07	73.41	**87.08**	**80.99**

The main purpose in this work is to monitor older people in an unstructured scene (e.g., home) and to recognize the types of activities they perform. We have extended the work in Ellomiietcv2014 that was basically an unsupervised method to learn behavioral patterns of individuals without restraining subjects to follow a predefined activity model. The main concern in previous work is to find different zones in the scene where activities take place (scene topology) by employing trajectory information provided by tracking algorithm. The previous work in Ellomiietcv2014 (**waiting hal acceptation**) proposes a Hierarchical Activity learning Model (HAM) to learn activities based on previously identified topologies. The current work examines the same potential while first, incorporating image descriptors [93] in a bag-of-word representation to differentiate actions in a supervised manner and second, combining the two approaches (supervised and unsupervised) to provide clues about actions inside each zone by classifying retrieved descriptors using a classifier.

Recently, dense trajectories are widely used for action recognition and have been shown state-of-the-art performance [93]. For the purpose of the current work, we use HOG and HOF descriptors for supervised action recognition. Figure 32 shows a general description of the supervised framework. For the learning phase, the dense trajectories are extracted from input images coming from RGBD camera. Following Ellomiietcv2014, three-level topology of the scene is constructed by trajectory information coming from tracking algorithm [62]. The topology is used to split input video stream into chunks by checking where the person is with respect to the learned zones. Then, for every video chunk, dense descriptors are extracted and stored. A codebook representation is obtained by applying a k-means clustering algorithm on the whole set of extracted features. Next, the action histograms are calculated by employing the codebook. A SVM classifier is trained and stored to use in test phase via calculated histograms.

Figure 32. Flow diagram for supervised action recognition.

In recognition phase, we similarly split the test videos by comparing each trajectory point with learned topologies, extract the descriptor for each split, and the histograms are calculated via k-NN using the codebook generated in learning phase. Then, the histograms are classified using the trained SVM classifier and resulting labels are evaluated by comparing with the ground truth.

We have assessed the performance of the supervised activity recognition framework using 183 video splits of 26 subjects. We divided the video dataset to training and testing groups. Training set includes 93 videos of 15 subjects and the test set includes 90 videos of 11 subjects. Notice that the number of videos is counted after splitting process has been done on input data. We used the videos recorded from CHU Nice hospital while real patients are visiting their doctors and are asked to perform several activities in specified locations of the room. The activities we considered in our tests include: “preparing tea”, “watching TV”, “using phone”, “reading on chair”, “using pharmacy”, and “using bus map”. For RGB-D camera, we have used the person detection algorithm in [79] and tracking algorithm in [62]. The classification results for using HOG and HOF descriptors and corresponding confusion matrices are depicted in Table 11 and in Table 12 . For SVM classifier, we used RBF kernel.

Table 11. Confusion matrix for recognition results for HOG descriptor

Activity Names	1	2	3	4	5	6
1 Watching TV	11	0	0	0	0	0
2 Preparing Tea	0	18	0	0	0	0
3 Reading in Chair	1	0	10	0	0	0
4 Using Bus Map	0	0	0	14	0	0
5 Using Pharmacy Basket	0	0	0	0	10	0
6 Using Phone	0	0	0	0	0	25
Total	98.89%					

Table 12. Confusion matrix for recognition results for HOF descriptor

Activity Names	1	2	3	4	5	6
1 Watching TV	4	1	1	4	0	1
2 Preparing Tea	0	5	0	5	0	8
3 Reading in Chair	1	0	2	4	0	4
4 Using Bus Map	0	0	0	13	0	1
5 Using Pharmacy Basket	1	0	0	0	9	0
6 Using Phone	0	1	1	5	0	19
Total	57.78%					

As a future work, we are going to benefit from the action descriptors to discriminate different activities occurring in the same zone.

6.24. Abnormal Event Detection in Videos and Group Behavior Analysis

Participants: Giuseppe Donatielo, Vania Bogorny, Serhan Cosar, Luis Campos Alvares, Carolina Garate, François Brémond.

keywords: activity recognition, abnormal events, group behavior analysis, trajectory clustering

This work addresses two different issues: (i) abnormal event detection and (ii) group behavior analysis in videos.

6.24.1. Abnormal Event Detection

For abnormal event detection we are proposing a fused approach that combines trajectory-based and pixel-based analysis. In this work we first discover the activity zones based on object trajectories, and we investigate abnormal events considering objects that move in wrong direction and/or with abnormal speed. Second, inside each zone we extract dense tracklets and using the clustering technique we discover different types of actions, and are able to distinguish between normal and abnormal actions inside each zone.

While existing approaches for abnormal behavior detection do either use trajectory based or pixel based methods, we propose a fused solution which can detect simple abnormal behavior based on speed and direction, as well as more complex behavior as abnormal activities. In a first step we automatically learn the zones of the scene where most activities occur, by taking as input the trajectories of detected mobiles, analyzing then statistical information of each mobile in each zone (speed and direction), through the use of a scale-resolution analysis. This approach implies a considerable complexity decrease of having huge data set and then an extensive impact of the algorithm speed, without losing useful information. Figure 33 shows an example of this first part.

Figure 33. Example of Trajectories (left), Trajectories over the grid (center), that represents a given scale resolution, Zones discovering (right)

The next step concerns a pixel based analysis inside each zone. This step takes as input each zone computed in the previous step and the bounding box of the object trajectories, and extracts action descriptors inside the bounding box of each object trajectory inside the zone. With this step we obtain the different body movements of each detected mobile inside a zone. By clustering the body motions and using Bags of Words, we detect different types of abnormal activities inside each zone. Figure 34 shows an example of what just mentioned.

The last step of our approach is a clustering operation of all information gathered in the previous two steps, that is for each mobile, speed, direction, and body movements-actions in each zone are applied to discriminate between different types of abnormal behavior in the scene. A flow diagram of our approach is presented in Figure 35 .

We have tested our approach on several real videos recorded. We show with experiments on two open datasets that our approach is able to detect several types of abnormal behavior.

6.24.2. Group Behavior Analysis

Group behavior analysis is focused on the extraction of groups based on object trajectories and the analysis is performed over the dense tracklets, computed for the groups bounding boxes. From the analysis of the dense tracklets we detect different levels of agitation. These works are ongoing and have not yet been published.

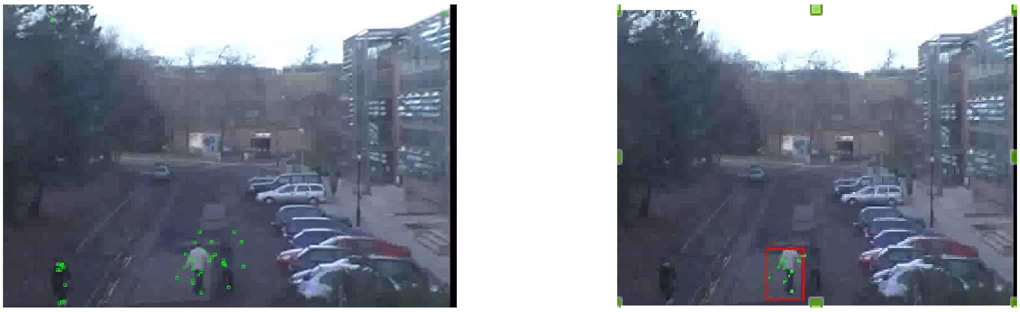


Figure 34. Tracked object

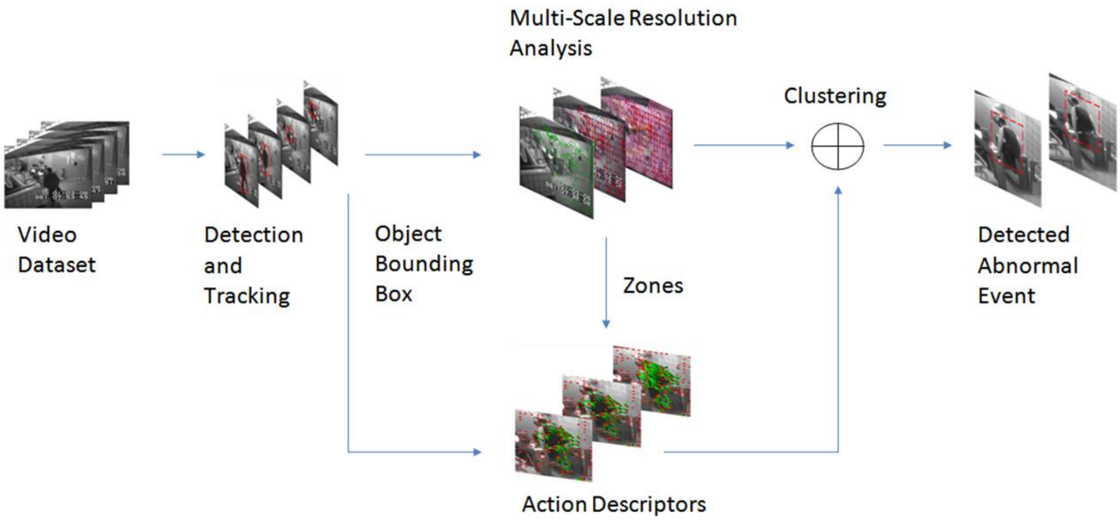


Figure 35. Flow Diagram of our approach

6.25. Model-Driven Engineering for Activity Recognition Systems

Participants: Sabine Moisan, Jean-Paul Rigault, Luis Emiliano Sanchez.

We continue to explore the applicability of model driven engineering (MDE) to activity recognition systems. Of course, setting up a complete methodology is a long term objective.

6.25.1. Feature Models

Features models are convenient representations of system variability but the drawback is a risk of combinatorial explosion of the number of possible configurations. Hence we have extended feature models with quality attributes and associated metrics to facilitate the choice of an optimal configuration, at deployment as well as at run time. We have proposed several strategies and heuristics offering different properties regarding optimality of results and execution efficiency [41].

This year we have conducted further experiments to evaluate the optimization algorithm and the metrics. In particular, we studied the prediction accuracy of the additive metrics for estimating two properties of interest: frame processing time and reconfiguration time. The goal was to compare predicted against measured properties of a running system. We used a simple video chain implemented with OpenCV components (acquisition, filtering, various detections, and visualisation) and we tested it on a video sample of 48s (i.e., about 1350 frames). We defined a feature model for this chain, which exhibits 14 valid configurations.

We first computed the properties of each component in isolation (based on a set of repetitive measurements), then we measured the actual frame processing time and reconfiguration time, and finally we compared the estimated and actual values.

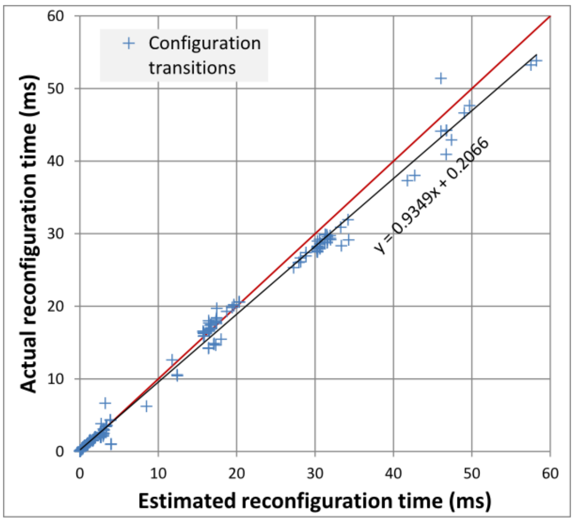


Figure 36. Measured and estimated reconfiguration time. Black line represents regression curve and red line the ideal trend

Figure 36 displays the results for reconfiguration time. In our experiment, we have a total of 210 transitions over a set of 15 valid configurations (14 system configuration + one for system shutdown). For frame processing time, we achieved an accuracy of 96.7% on average, and for reconfiguration time the accuracy was between 90.5% and 87.6%.

Introducing metrics in feature models is a precious help to reduce the huge set of valid configurations after a dynamic context change and to provide a real time selection of an appropriate running configuration. However, more evaluation remains to be done on other runtime properties and for other video chains.

6.25.2. Configuration Adaptation at Run Time

To react to environment changes we favor the “model at run-time” approach. Our current prototype ressorts to tools and technologies which were readily available. This made possible a proof of concepts. However, this induces several redundant representations of the same data, consistency problems, coexistence of several formalisms or programming languages, and superfluous back and forth inter-module communications. This year we started to design a more homogeneous and better integrated prototype. The two key points are, first, a component management framework, second, a feature model managemen tool.

This year, we have continued to develop our OSGi-like component framework, but more adapted to real time and compatible with our extended feature models. Concerning feature model management, we started to study a FAMILIAR replacement that integrates smoothly into the C++ video analysis chain instead of being a separate (Java) tool. Moreover, the new tool should cope with our feature extensions (e.g., quality attributes).

6.26. Scenario Analysis Module

Participants: Annie Ressouche, Sabine Moisan, Jean-Paul Rigault, Daniel Gaffé, Omar Abdalla.

Keywords: Synchronous Modelling, Model checking, Mealy machine, Cognitive systems.

To generate activity recognition systems we supply a scenario analysis module (SAM) to express and recognize complex events from primitive events generated by SUP or other sensors. The purpose of this research axis is to offer a generic tool to express and recognize activities. Genericity means that the tool should accommodate any kind of activities and be easily specialized for a particular framework. In practice, we propose a concrete language to specify activities in the form of a set of scenarios with temporal constraints between scenarios. This language allows domain experts to describe their own scenario models. To recognize instances of these models, we consider the activity descriptions as synchronous reactive systems [80] and we adapt usual techniques of synchronous modelling approach to express scenario behaviours. This approach facilitates scenario validation and allows us to generate a recognizer for each scenario model.

Setting up our tools on top of an existing language such as LUSTRE was convenient for rapid prototyping. However, it appeared delicate for efficiency reasons on the one hand, but also because it is a closed environment, difficult to customize. Hence we developed our own language LE and its environment CLEM (see section 6.27). This year, we focus on the expression of scenario models in CLEM through the internship of Omar Adballa [51] and we define in CLEM a specific back end to generate recognition engines (see figure 37). However, mastering all aspects of this environment will allow the user scenario description language to rely directly on the semantics of LE and not on its syntax. This reduces the number of necessary translations.

Currently, SAM implements an “exact” algorithm in the sense that it generates, at each instant, *all* possible scenario instances although many of them will freeze, still holding system resources. We have started scalability studies to evaluate the risk of combinatorial explosion. In parallel we enriched the synchronous scenario descriptions to reduce the number of generated scenario instances as well as the number of instances to awake at each instant. We are currently modifying our recognition engine generator to take advantage of this supplementaty information.

6.27. The Clem Workflow

Participants: Annie Ressouche, Daniel Gaffé, Mohamed Bouatira, Ines Sarray.

Keywords: Synchronous languages, Synchronous Modelling, Model checking, Mealy machine.

This research axis concerns the theoretical study of a synchronous language LE with modular compilation and the development of a toolkit (see Figure 37) around the language to design, simulate, verify and generate code for programs. The novelty of the approach is the ability to manage both modularity and causality. This year, we focus on the improvement of both LE language and compiler concerning data handling and in the generation of back-ends required by other research axis of the team (see 6.26 and 6.28). We also improve the design of a new simulator for LE programs which integrates our new approach.

First, synchronous language semantics usually characterizes each output and local signal status (as present or absent) according to input signal status. To reach our goal, we defined a semantics that translates LE programs into equation systems. This semantics bears and grows richer the knowledge about signals and is never in contradiction with previous deduction (this property is called constructiveness). In such an approach, causality turns out to be a scheduling evaluation problem. We need to determine all the partial orders of equation systems and to compute them, we consider a 4-valued algebra to characterize the knowledge of signal status (unknown, present, absent, overknown). Last year, we chose an algebra which is a bilattice and we showed that it is well suited to solve our problem. To compute the partial orders of equation systems, we introduced two ways : a CPM (Critical Path Method) like algorithm, efficient and a "fix point" approach that allows us to show that we can compute partial orders locally and in an incremental way, or globally (thanks to the uniqueness if fix points). We introduced "abstract" equation systems and our method allows us to compute their partial orders. We defined a new intermediate format LEA (see figure 37) to record these sorted abstract eaquation systems and they will be expanded latter in concrete equation systems with a refinement operation. We apply this technique to the compilation of Grafcet language. Our work in under publication in two journal papers.

In CLEM, we added types and data a few years ago and this year we complete this addition, we know can express automata where control and output signals are valued. From last year, we rely on CLEM both to design SAM (see section 6.26) and to perform validation in a component-based middleware (see section 6.28). To this aim, we generate now two specific output formats dedicated to these applications [54]

Finally, in CLEM, we generate an independent intermediate code (LEC) before specific target generations. This code represents the semantics of programs with 4-valued equation systems. In our design flow, we need to simulate programs at this level. This year, we complete a simulator begun last year but which did not integrate the data part of the language. The simulator GUI has been designed again in Qt and the simulator takes into account the values carried by signals. This work has been done by Mohamed Bouatira during his internship.

6.28. Multiple Services for Device Adaptive Platform for Scenario Recognition

Participants: Annie Ressouche, Daniel Gaffé, Ines Sarray, Jean-Yves Tigli.

Keywords: Synchronous Modelling, Model checking, Mealy machine, Ubiquitous Computing.

The aim of this research axis is to federate the inherent constraints of an activity recognition platform like SUP (see section 5.1) with a service oriented middleware approach dealing with dynamic evolutions of system infrastructure. The Rainbow team (Nice-Sophia Antipolis University) proposes a component-based adaptive middleware (WComp [91], [90], [73]) to dynamically adapt and recompose assemblies of components. These operations must obey the "usage contract" of components. The existing approaches don't really ensure that this usage contract is not violated during application design. Only a formal analysis of the component behaviour models associated with a well sound modelling of composition operation may guarantee the respect of the usage contract.

The approach we adopted introduces in a main assembly, a synchronous component for each sub assembly connected with a critical component. This additional component implements a behavioral model of the critical component and model checking techniques apply to verify safety properties concerning this critical component. Thus, we consider that the critical component is validated.

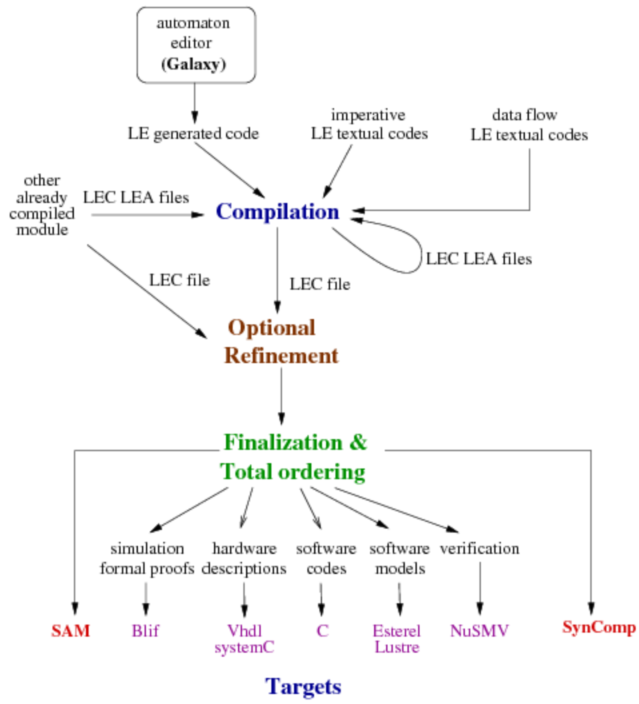


Figure 37. The CLEM *toolkit*

In [83], [82], we showed that an efficient means to define the synchronous components which allow to validate critical component behaviours, is to specify them with Mealy machines. Previously, we used a classical synchronous language (Lustre) to specify synchronous components, but the integration of the synchronous component code into WComp was not straightforward because Lustre compiler is not opened and cannot integrate new target code needed by WComp. This year, we rely on CLEM to design synchronous monitor as CLEM automata and we extend CLEM to generate automatically the internal code of WComp(C#).

On another hand, for some critical components, we can be led to introduce several synchronous monitors, each of them being related to a sub assembly. Then, some outputs of these synchronous monitors can be linked to the same input of a critical component. To face this problem, we introduced in [83], [82] a composition under constraints operation, which composed all the synchronous monitors linked to a critical component according to a set of constraints telling us how the respective outputs of monitors which access the same input are combined. We proved that this operation preserved safety properties, however it cannot ensure adaptivity and incrementality. This year, we have introduced a new way to compose synchronous monitors. We introduce an additional monitor (called constrain monitor) which express as a CLEM Mealy machine (set of equations) the constraints and we perform a usual parallel composition of all the monitors. Moreover, we define a syntactic mean (language DCL) to describe constraints in a generic way and we can derive automatically the constraint monitor for each specific application. In such a setting, we are able to manage the appearance and the desappearance of synchronous monitors.

Moreover, in middleware as WComp, communication is asynchronous while synchronous monitors respect the synchrony paradigm. Thus, we need to introduce in the middleware a means to transform asynchronous events into synchronous entries for synchronous monitors and the opposite to connect again the synchronous events into asynchronous design. To this aim, we introduced in WComp, specific components that receive asynchronous events and generate synchronous ones. Indeed, the part of this component is to decide how asynchronous events will be gather into a synchronous instant (logical time). To this aim, the user can choose between different politics. Then, all the events composing an instant are serialized and deserialized by the synchronous monitor. The desynchronisation operation performs the opposite way[54].

STEEP Team

6. New Results

6.1. Highlights of the Year

This year has seen a number of major advances in the team research projects, on several fronts. The first one concerns the most important and time consuming project, namely integrated land use, activity and transport modelling (LUTI modelling). In this respect, the results described in 6.8 below constitute probably the first set of works contributing sophisticated numerical procedures to the calibration and validation of the TRANUS LUTI model.

The second significant breakthrough concerns the completion of a downscaling methof for Material Flow Analysis (MFA), a key aspect in the characterization and understanding of territorial metabolism for decision-help purposes (section 6.2).

Finally, the modelling effort on land use change for the ESNET project has now been mostly completed, and an operational LUCC model has been calibrated and validated for this project (section 6.3).

6.2. Downscaling Material Flow Analysis: the case of the cereals supply chain in France

The spatial reconstruction of the production, trade, transformation and consumption flows of a specific material can become an important decision-help tool for improving resource management and for studying environmental pressures from the producer's to the consumer's viewpoint. One of the obstacles preventing its actual use in the decision-making process is that building such studies at various geographical scales proves to be costly both in time and manpower. We propose a semi-automatic methodology to overcome this issue. First a supply chain model at the national level has to be designed. Supply and use tables are used to handle the data consistently. Finding the appropriate level of detail for both products and industries is an iterative process: with a small number of highly aggregated product categories, the study isn't likely to provide useful information while with a very detailed list of products and industries, finding input data, especially at local scales, won't be feasible. Secondly, national production, transformation, trade and consumption data have to be reconciled in order to respect the law of mass conservation: this is done through constraint optimization. Thirdly, regional supply and use tables are generated (either with direct data or through downscaling of national data using local proxies, e.g. employment statistics) and reconciled, taking into account the additional constraint that regional data must add up to national one.

We applied the methodology on the case of cereals and reconstructed the supply chain flows of the 22 French regions as well as the flows of four nested territories: France, the Rhône-Alpes région, the Isère département and the territory of the SCOT of Grenoble. Uncertainties of output data were estimated via Monte-Carlo simulations. We display the results using our Sankey diagram vizualisation tool. A research paper is in the reviewing process for one of the major journals in this field.

Future steps include coupling this model with economic (added value), social (local employment) and environmental (environmental pressures) aspects in order to provide new information to decision-makers at various administrative levels (from a group of cities to the national level).

6.3. Mapping and land use and land cover change for the ESNET project

The ESNET project (EcoSystem services NETworks) is a collaboration lead by LECA (Laboratoire d'ECologie Alpine, UJF) that aims at characterizing the ecosystem services of the Grenoble urban region (about 2/3 of the Isere département) at the 2040 horizon under various constraints of urban policy planning, changes in agricultural and forest management, and climate change impact on ecosystems. A preliminary task in this research program was the elaboration of very detailed maps (both in terms of land use and of resolution) of the study area at three different dates (1998, 2003 and 2009) based on available satellite and IGN data, in order to characterize past land use patterns as well as agricultural rotation patterns. These have been made and completed at Inria with the hiring of specialized engineers in these tasks, funded by the ESNET program. This exercice informs the next task (land use and land cover change – LUCC – modelling). Hosting this work at Inria was not only logical in terms of the available computer environment, but also useful in terms of visibility of Inria from outside planning agencies.

The LUCC model itself is developed partly at Inria (for modelling expertise) and partly at LECA (for expertise on ecological change drivers). The model development is now operational, thanks to a major effort on this front in 2014. Both transitions from non urban to urban and use and changes of agricultural practices are now calibrated and validated. The first scenario has been successfully simulated in terms of land use. The three other scenairios of the project are in the final stage of elaboration before simulation, so that the land use change simulation phase of the ESNET project should be completed by the end of April, 2015.

Two resarch papers are in the process of being written on the question of land use practices and their evolution in the study area, and a third one on issues of principle in land use modelling is also underway.

6.4. Benchmarking tools for the climate negotiation of GHG emission reduction trajectories

Climate negotiations related to global warming are another important issue of sustainable development. In this framework that is place at international scale we propose a benchmarking tool that is designed to avoid the main limitations of actual negotiation schemes. Our approach is based on the original Soft Landing proposition, made by Criqui and Kouvaritakis in the early 2000. We develop an up to date solution which improves the original idea mainly by introducing common but differentiated emission reduction profiles and by developing a dedicated algorithm for that purpose (called REDEM). To be compatible with global objectives, it is commonly accepted that for most developing regions, the national emission curves should admit a maximum and then should progressively decline. Similarly, we emphasize the fact that, in order to achieve the global objectives, all states will have to entail mitigation efforts, the intensity which may be measured by the rate of variation of the national emissions. At one point, the effort will reach a maximum, when the rate of variation in absolute value is at its maximum, and then decrease. In other words, there will also be a peak in the effort. Then we propose to base the benchmark on this peak of effort. This work has been done in collaboration with EDDEN Laboratory, in particular Patrick Criqui and Constantin Ilasca. It has been published in [5].

6.5. On the acceptability of land use transport integrated models by French end users as operational tools: from understanding to daily use

Land Use and Transport Integrated models (LUTIs) are promising approaches for urban planning. There is large literature describing their technical architectures or using them in various scientific contexts. Yet little attention has been paid to expectations of practitioners (planners) and to the daily use of such models. There is clearly an important gap between research and practice: a daily use of LUTIs for the simulation of regional planning policies is still an exception in France, despite important research investments and recent interest of planning agencies., and this situation does not seem to be specific to France. We worked on sheding light on what would make them definitely accepted and more used by planners to evaluate a range of urban and transport policies. To do so, we have interviewed different types of end users in France to identify their motivations and barriers to use LUTI models, in addition to literature study and our own experience dealing with urban planning agencies. We have analysed the main obstacles that prevent LUTIs from being widely used by local authorities. It is important to identify that there are two main issues: 1) Do current LUTIs really answer the questions and practical issues territorial agencies are confronted with on a day-to-day basis? Do they match their interests and expectations? 2) Are current LUTIs suitable with respect to the constraints and limitations of local agencies? The main obstacles associated with these issues are: first, it is difficult to match rather generic models with very specific and varied end users questions; second, it is costly and heavy to implement and use a LUTI (capacity obstacles); third, there is no guarantee that results of a dedicated LUTI will have any impact on the policy design (decision making obstacles). The results of our analysis show demand for a far more bottom-up oriented approach: the models should consider objectives and general needs of end users to live up to their expectations. Only a closer collaboration between modelers and end users, and more efforts to integrate modeling into urban planning, will make LUTIs considered as relevant approaches.

This work has been done in collaboration with Mathieu Saujot (IDDRI) and Mathieu De Lapparent (IFSSTAR), and belongs to the work program of CiTIES project.

6.6. Replication procedure for grouped Sobol' indices estimation in dependent uncertainty spaces

Sensitivity analysis studies how the uncertainty on an output of a mathematical model can be attributed to sources of uncertainty among the inputs. Global sensitivity analysis of complex and expensive mathematical models is a common practice to identify influent inputs and detect the potential interactions between them. Among the large number of available approaches, the variance-based method introduced by Sobol' allows to calculate sensitivity indices called Sobol' indices. Each index gives an estimation of the influence of an individual input or a group of inputs. These indices give an estimation of how the output uncertainty can be apportioned to the uncertainty in the inputs. One can distinguish first-order indices that estimate the main effect from each input or group of inputs from higher-order indices that estimate the corresponding order of interactions between inputs. This estimation procedure requires a significant number of model runs, number that has a polynomial growth rate with respect to the input space dimension. This cost can be prohibitive for time consuming models and only a few number of runs is not enough to retrieve accurate informations about the model inputs.

The use of replicated designs to estimate first-order Sobol' indices has the major advantage of reducing drastically the estimation cost as the number of runs becomes independent of the input space dimension. The generalization to closed second-order Sobol' indices relies on the replication of randomized orthogonal arrays. The motivation of this work is to extend this methodology in presence of dependent inputs. Indeed, the case of correlated parameters has to be tackled with caution, as the calculation of single input indices does not provide anymore a proper information, that can be easily interpreted. One strategy is thus to define grouped indices for groups of correlated variables. We address this issue by proposing an approach based on replicated designs and randomized orthogonal arrays that enables to take into account dependency within inputs. We suppose that this dependency can be expressed through constraints. This approach can be used facing any set of constraints at the condition that one is able to provide points in the input space that verify the considered constraints. Guided by our application on a land-use and transport integrated model (LUTI) where some economical parameters are linked by order relations, we focus on the case of sets of linear ordered constraints. Thus we propose a sampling strategy based on the simplex geometric structure, that ensures a proper input space filling.

This work has been done in collaboration with Laurent Gilquin and Clementine Prieur (members of Moise Team), and belongs to the work program of CiTIES project. It is described in [18]. The proposed procedure will be soon applied to study the sensitivity of TRANUS model.

6.7. Specifications for the calibration of Simbad model

"Simbad" is a LUTI model developed by LET. In the context of the CITiESANR project, we have done a comprehensive and detailed study of the parameters of the model in order to fully specify the calibration process of the model. For example, we have specified the objects of interest and indicators, as well as satisfaction criteria. This work has been done in close collaboration with LET.

6.8. Calibration of the TRANUS Land Use Module

The setting up of a LUTI model requires, like most numerical models, at least one phase of parameter estimation. This is concisely referred to here as calibration, although the calibration of a LUTI model also entails other aspects such as the definition of spatial zones, of economic sectors, etc. The TRANUS LUTI model plus software, like many other existing models, come along with a relatively simple calibration methodology. Most LUTI models indeed perform parameter estimation in a piecewise fashion, by sequentially estimating subsets of parameters. While this reduces the mathematical and computational complexity of calibration, neglecting the interactions across different modules and their parameters, may result in a significant loss of a model's quality. A second issue is that TRANUS, like several other LUTI softwares, employs rudimentary numerical routines for parameter estimation. We aim at reducing these weaknesses.

To do so, we first defined a particular parameter estimation problem for TRANUS properly as an optimisation problem, based on an explicit cost function that is to be minimised (something lacking in many articles on LUTI calibration). Next, we developed a series of numerical estimation schemes to solve this optimisation problem. The main difficulty here was that the model is dynamic; by delving into the model's equations and structure, we were able to unwind the model's dynamics and to make it amenable to standard numerical optimisation by gradient descent type methods [4]. This was first done for the estimation of a particular subset of model parameters (the so-called shadow prices). We have recently started to work on the simultaneous estimation of these and other model parameters.

This work is done in collaboration with Arthur Vidard from the MOISE Inria project-team and Brian Morton from the University of North Carolina at Chapel Hill.

6.9. State of the Art on the Calibration and Validation of LUTI Models

One of the tasks of the CITiES project is to construct an extensive state of the art report on the calibration and validation of LUTI models. We coordinate this effort, which involves all partners of CITiES, together with the project partner LVMT (Nicolas Coulombel). It consists of the definition of a taxonomy, of an extensive literature research and of a critical analysis of this literature. A short publication that explains the goals of this effort and some intermediate findings, has been presented in [3]. The completion of this task is expected for the first semester of 2015.

SUMO Project-Team

6. New Results

6.1. Highlights of the Year

We started our first industrial collaboration "Project P22" with Alstom Transport, in the context of a common laboratory between Inria and Alstom. The project started in March 2014 and tackles robustness issues and regulation in urban train systems. The second phase of the project will start in march 2015, for a duration of three years. Most of the researchers of Sumo are involved in this project.

6.2. Control and enforcement

6.2.1. Runtime enforcement of timed properties

Participants: Thierry Jéron, Hervé Marchand, Srinivas Pinisetty.

Runtime enforcement is a powerful technique to ensure that a running system satisfies some desired properties. Using an enforcement monitor, an (untrustworthy) input execution (in the form of a sequence of events) is modified into an output sequence that complies with a property. Over the last decade, runtime enforcement has been mainly studied in the context of untimed properties. The contributions [26] and [34] deal with runtime enforcement of timed properties by revisiting the founda-tions of runtime enforcement when time between events matters. We propose a new enforcement paradigm where enforcement mechanisms are time retardants: to produce a correct output sequence, additional delays are introduced between the events of the input sequence. We consider runtime enforcement of any regular timed property defined by a timed automaton. We prove the correctness of enforcement mechanisms and prove that they enjoy two usually expected features, revisited here in the context of timed properties. The first one is soundness meaning that the output sequences (eventually) satisfy the required property. The second one is transparency, meaning that input sequences are modified in a minimal way. We also introduce two new features, i) physical constraints that describe how a time retardant is physically constrained when delaying a sequence of timed events, and ii) optimality, meaning that output sequences are produced as soon as possible. To facilitate the adoption and implementation of enforcement mechanisms, we describe them at several complementary abstraction levels. Our enforcement mechanisms have been implemented and our experimental results demonstrate the feasibility of runtime enforcement in a timed context and the effectiveness of the mechanisms. Finally, in [33], we considered more practical applications. Indeed, in network security, RE monitors can detect and prevent Denial-of-Service attacks. In resource allocation, RE monitors can ensure fairness. Specifications in these domains express data-constraints over the received events where the timing between events matters. To formalize these requirements, we introduce Parameterized Timed Automata with Variables (PTAVs), an extension of Timed Automata (TAs) with internal and external variables. We then extend enforcement for TAs to enforcement for PTAVs for safety properties. We model requirements from the considered application domains and show how enforcement monitors can ensure system correctness w.r.t. these requirements.

6.2.2. Enforcing opacity

Participant: Hervé Marchand.

In [22], we have been interested in enforcing opacity of regular predicates on modal transition systems. Intuitively, a labelled transition system $\mathcal{T}$ partially observed by an attacker, and a regular predicate S over the runs of $\mathcal{T}$, enforcing opacity of the secret S in $\mathcal{T}$ means computing a supervisory controller K such that an attacker who observes a run of the controlled system $K \setminus \mathcal{T}$ cannot ascertain that the trace of this run belongs to S based on the knowledge of $\mathcal{T}$ and K. We lift the problem from a single labelled transition system $\mathcal{T}$ to the class of all labelled transition systems specified by a *Modal Transition System*$\mathcal{M}$. The lifted problem is to compute the maximally permissive controller K such that S is opaque in $K/\mathcal{T}$ for every labelled transition system $\mathcal{T}$ which is a model of $\mathcal{M}$. The situations of the attacker and of the controller are asymmetric: at run time, the attacker may fully know $\mathcal{T}$ and K whereas the controller knows only $\mathcal{M}$ and the sequence of actions executed so far by the unknown $\mathcal{T}$.

In [23], we provided a different solution by enforcing and validate ay runtime various notion of opacity. More specically, we studied how we can model-check, verify and enforce at system runtime, several levels of opacity. Besides existing notions of opacity, we also introduce K-step strong opacity, a more practical notion of opacity that provides a stronger level of confidentiality.

6.2.3. Discrete Controller Synthesis for Infinite State Systems with ReaX

Participants: Nicolas Berthier, Hervé Marchand.

This year, we investigated the control of infinite reactive synchronous systems modeled by arithmetic symbolic transition systems for safety properties handling numerical variable. We provide effective algorithms allowing to solve the safety control problem, and report on experiments based on ReaX, our tool implementing these algorithms [28].

6.3. Model expressivity and quantitative verification

6.3.1. Diagnosis

Participants: Nathalie Bertrand, Sébastien Chédor, Éric Fabre, Loïc Hélouët, Blaise Genest, Hervé Marchand, Christophe Morvan.

Diagnosis of a system consists in providing explanations to a supervisor from a partial observation of the system and a model of possible executions. This year, we have extended results on diagnosis algorithm from scenarios. Systems are modeled using High-level Message Sequence Charts (HMSCs), and the diagnosis is given as a new HMSC, which behaviors are all explanations of the partial observation. The results published this year are first an offline centralized diagnosis algorithm (a single process in a network collects an observation, and emits a diagnosis) that has then been extended to a decentralized version of this algorithm. This allows us to give a complete diagnosis framework for infinite state systems, with a strong emphasis on concurrency and causal ordering in behaviors. HMSC-based diagnosis showed nice properties w.r.t. compositionality. We have also considered solutions for online diagnosis from scenarios, but came to the conclusion that online solutions are memory consuming, and need too many restrictions to run with finite memory. The last contribution of this work is an application of diagnosis techniques to anomaly detection, that is a comparison of observation of the system with a model of usual behaviors to detect security attacks. This work has been published this year [24].

In [21] we have been interested in the analysis of discrete event systems under partial observation which is an important topic, with major applications such as the detection of information flow and the diagnosis of faulty behaviors. These questions have, mostly, not been addressed for classical models of recursive systems, such as pushdown systems and recursive state machines. In this paper, we consider recursive tile systems, which are recursive infinite systems generated by a finite collection of finite tiles, a simplified variant of deterministic graph grammars (slightly more general than pushdown systems). Since these systems are infinite-state in general powerset constructions for monitoring do not always apply. We exhibit computable conditions on recursive tile systems and present non-trivial constructions that yield effective computation of the monitors.We apply these results to the classic problems of state-based opacity and diagnosability (off-line verification of opacity and diagnosability, and also run-time monitoring of these properties). For a decidable subclass of recursive tile systems, we also establish the decidability of the problems of state-based opacity and diagnosability.

In discrete event systems prone to unobservable faults, a diagnoser must eventually detect fault occurrences. The diagnosability problem consists in deciding whether such a diagnoser exists. We laid the foundations of diagnosis and predicatability for probabilistic systems represented by partially observed Markov chains (denoted pLTS) [32]. In particular, we studied different specifications of diagnosability and establish their relations both in finite and infinite pLTS. Then we analyzed the complexity of the diagnosability problem for finite pLTS: we showed that the polynomial time procedure proposed earlier is erroneous and that in fact for all considered specifications, the problem is PSPACE-complete. We also established tight bounds for the size of diagnosers. Afterwards we considered the dual notion of predictability which consists in predicting that in a safe run, fault will eventually occur. Predictability is easier than diagnosability: it is NLOGSPACE-complete. Yet the predictor synthesis is as hard as the diagnoser synthesis.

When a system is not diagnosable, the active diagnosis problem consists in controlling the system in order to ensure its diagnosability. In the same probabilistic setting, the active diagnosis problem consists in deciding whether there exists some observation-based strategy that makes the system diagnosable with probability one. We proved that this problem is EXPTIME-complete, and that the active diagnosis strategies are belief-based. The *safe* active diagnosis problem is similar, but aims at enforcing diagnosability while preserving a positive probability to non faulty runs, i.e. without enforcing the occurrence of a fault. We prove that this problem requires non belief-based strategies, and that it is undecidable. However, it belongs to NEXPTIME when restricted to belief-based strategies. Our work also refines the decidability/undecidability frontier for verification problems on partially observed Markov decision processes [30].

6.3.2. Probabilistic model checking

Participants: Nathalie Bertrand, Blaise Genest, Paulin Fournier.

In [16], we considered the verification of Markov chains against properties talking about distributions of probabilities. Even though a Markov chain is a very simple formalism, by discretizing in a finite number of classes the space of distributions through some symbolics, we proved that the language of trajectories of distribution (one for each initial distribution) is not regular in general, even with 3 states. We then proposed a parametrized algorithm which approximate what happens to infinity, such that each symbolic block in the approximate language is at most ϵ away from the concrete distribution.

Parameterized verification aims at validating a model of a system irrespective of the value of a parameter. This year, we studied verification problems for a model of network with the following characteristics: the number of entities is parametric, communication is performed through broadcast with adjacent neighbors, entities can change their internal state probabilistically and reconfiguration of the communication topology can happen at any time. The semantics of such a model is given in term of an infinite state system with both non deterministic and probabilistic choices. We are interested in qualitative problems like whether there exists an initial topology and a resolution of the non determinism such that a configuration exhibiting an error state is almost surely reached. We showed in [44] that all the qualitative reachability problems are decidable and some proofs are based on solving a 2 player game played on the graphs of a reconfigurable network with broadcast with parity and safety objectives.

On a different topic, we considered a control problem for stochastic sytems specified by timed automata with distributions over delays. In [29] we considered reachability objectives on such decision stochastic timed automata (DSTA). Given a reachability objective, the value 1 problem asks whether a target can be reached with probability arbitrarily close to 1. Simple examples show that the value can be 1 and yet no strategy ensures reaching the target with probability 1. In this paper, we prove that, the value 1 problem is decidable for single clock DSTA by non-trivial reduction to a simple almost-sure reachability problem on a finite Markov decision process. The ε-optimal strategies are involved: the precise probability distributions, even if they do not change the winning nature of a state, impact the timings at which ε-optimal strategies must change their decisions, and more surprisingly these timings cannot be chosen uniformly over the set of regions.

6.3.3. Distributed timed systems

Participants: Blaise Genest, Loïc Hélouët.

We have proposed and considered properties of a new timed variant of Petri nets [42], namely Timed Petri Nets with Urgency, that extend Timed Petri Nets with the main features of TPNs. Time Petri Nets (TPN) [52] and Timed Petri Nets [45] are two incomparable classes of concurrent models with timing constraints: urgency cannot be expressed using Timed Petri Nets, while TPNs can only keep track of a bounded number of continuous values (clocks). The work performed this year provides up to-our-knowledge the first decidability results for Petri Net variants combining time, urgency and unbounded places. We have obtained decidability of control-state reachability for the subclass of Timed Petri Nets with Urgency where urgency constraints can only be used on bounded places. By restricting this class to use a finite number of clocks, we have shows decidability of (marking) reachability. Formally, this class corresponds to TPNs under a new, yet natural, timed semantics where urgency constraints are restricted to bounded places. Further, under their original semantics, reachability for a more restricted class of TPNs is decidable.

6.3.4. Test Generation from Recursive Tile Systems

Participants: Sébastien Chédor, Christophe Morvan, Thierry Jéron.

In [20] we explore the generation of conformance test cases for *Recursive Tile Systems* in the framework of the classical **ioco** testing theory. The RTS model allows the description of reactive systems with recursion, and is very similar to other models like Pushdown Automata, Hyperedge Replacement Grammars or Recursive State Machines. Test generation for this kind of infinite state labelled transition systems is seldom explored in the literature. The first part presents an off-line test generation algorithm for *Weighted* RTSs, a determinizable subclass of RTSs, and the second one, an on-line test generation algorithm for the full RTS model. Both algorithms use test purposes to guide test selection through targeted behaviours. Additionally, essential properties relating verdicts produced by generated test cases with both the soundness with respect to the specification, and the precision with respect to a test purpose, are proved.

6.4. Management of large distributed systems

6.4.1. Distributed optimal planning

Participant: Éric Fabre.

Planning problems consist in organizing actions in a system in order to reach one of some target states. The actions consume and produce resources, can of course take place concurrently, and may have costs. We have a collection of results addressing this problem in the setting of distributed systems. This takes the shape of a network of components, each one holding private actions operating over its own resources, and shared/synchronized actions that can only occur in agreement with its neighbors. The goal is to design in a distributed manner a tuple of local plans, one per component, such that their combination forms a consistent global plan of minimal cost.

Our previous solutions to this problem modeled components as weighted automata. In collaboration with Loig Jezequel (TU Munich) and Victor Khomenko (Univ. of Newcastle), we have extended this approach to the case of components modeled as safe Petri nets [50]. This allows one to benefit from the internal concurrency of actions within a component. Benchmarks have shown that this method can lead to significant time reductions to find feasible plans, in good cases. In the least favorable cases, performances are comparable to those obtained with components modeled as automata. The method does not apply to all situations however, as computations require to perform ϵ-reductions on Petri-nets (our work also contains a contribution to this difficult question). This work has been accepted by the ACM Transactions in Embedded Computing Systems, to appear in 2015.

6.5. Data driven systems

6.5.1. Web services

Participants: Blaise Genest, Loïc Hélouët.

This year, we considered transactional properties (ACID) for web services. In particular, we focused on the atomicity (A of ACID) property, obtained in case of a failure inside an atomic block through compensation of the executed actions of the block. To do so, logs need to be kept. We were interested in maintaining the maximal amount of privacy. We proposed modular algorithms [19] which maintain privacy between modules, with minimal information shared among modules, both in the logging and the compensation phases. Furthermore, each module logs a small number of information, such that the sum of all actions logged is guaranteed minimal. Last, modularity allows fast algorithms, as they need to consider only what happens in the module itself, and not the exact structure of its parent module nor of its sub-modules.

We also published results on our model of sessions systems [27]. This models allows for the modeling of distributed web-based systems that are running an arbitrary number of transactions among arbitrarily many participants. We have shown how simple restrictions can guarantee decidability of simple coverability properties, and then be used to detect violation of buisness rules such as conflict of interest, or a more complex property called the chinese wall.

We are currently considering new models that manage at the same time explicit workflows and structured data. This model can be seen as a combination of AXML [46] and Petri nets.

6.5.2. An Artifact-centric Process Model

Participants: Éric Badouel, Loïc Hélouët, Christophe Morvan.

In [37] we present a purely declarative approach to artifact-centric case management systems, and a decentralization scheme for this model. Each case is presented as a tree-like structure; nodes bear information that combines data and computations. Each node belongs to a given stakeholder, and semantic rules govern the evolution of the tree structure, as well as how data values derive from information stemming from the context of the node. Stakeholders communicate through asynchronous message passing without shared memory, enabling convenient distribution.

TACOMA Team

6. New Results

6.1. Self-describing objects and tangible data structures

Participants: Nebil Ben Mabrouk, Paul Couderc [contact], Arnab Sinha.

A development in the line of the coupled objects principles are self-describing objects. While previous works enabled integrity checking over a set of physical objects, these mechanisms were limited in two aspects: expressiveness and autonomy. More precisely, coupled objects support the detection of special conditions (such as a missing element), but not the characterization of these conditions (such as describing the problem, identifying the missing element). Moreover, this compromises the autonomous feature of coupled objects, which would depend on external systems for analyzing these special conditions. Self-describing objects are an attempt to overcome these limitations, and to broaden the application perspectives of autonomous RFID systems.

The principle is to implement distributed data structure over a set of RFID tags, enabling a complex object (made of various parts) or a set of objects belonging to a given logical group to "ÃÂÃÂself-describe" itself and the relation between the various physical elements. Some applications examples includes waste management, assembling and repair assistance, prevention of hazards in situations where various products / materials are combined etc. The key property of self-describing objects is, like for coupled objects, that the vital data are self-hosted by the physical element themselves (typically in RFID chips), not an external infrastructure like most RFID systems. This property provides the same advantages as in coupled objects, namely high scalability, easy deployment (no interoperability dependence/interference), and limited risk for privacy.

However, given the extreme storage limitation of RFID chips, designing such systems is difficult:

- Data structures must be very frugal in terms of space requirements, both for the structure and for the coding.
- Data structures must be robust and able to survive missing or corrupted elements if we want to ensure the self-describing property for a damaged or incorrect object.

In the context of RFID system, the resiliency property of such data structures enables new information architecture and autonomous (offline) operation, which is very important for some RFID applications. On this topic, a generic graph structure applicable to RFID systems for supporting self-describing objects is proposed in Arnab Sinha's thesis document [1], and was published in [4].

6.2. Pervasive support for RFIDs

Participants: Nebil Ben Mabrouk, Paul Couderc [contact].

In situations where we have to read large collection of objects of various types, the performance is difficult to predict but may still be adequate for a given application. For example, some application can tolerate missing some tags, provided that miss read probability could be characterized. In some cases, read reliability could be improved using mechanical approaches, such as introducing movements in objects or antenna to introduce radio diversity during read. Finally, distributed data structure can be used over a set of tags to be used to mitigate the impact of mis-read (by using data redundancy) and to help the reading protocol by integrating hints about the tag set collection being read.

Our objective here is to study extensively by experimentation the behaviour of existing RFID solutions in the context of uncontrolled environment (meaning, random placement of tags on objects mixing various materials) in order to characterize their real-world performance regarding the parameters of such as tags numbers, density, frequencies, reader antenna design, dynamicity of objects (movements), etc. From these experimentations, we would like to identify the conditions that are favorable to acceptable performance, and the way where there are hopes of improvement with specific design for these difficult environments. These results should also allow improving the performance: high level integrity checks can guide low level operations by determining whether inventories are complete or not. This cross layer strategy can enable faster are more efficient inventory protocols.

An important milestone was completed in 2014, with the implementation of an experiment test bed in order to support the experiment campaign. This task involved a significant development and engineering effort. This testbed is currently deployed at the IETR (http://www.ietr.fr) building, and features a multi-axis mobile RFID antenna system driven by a software platform.

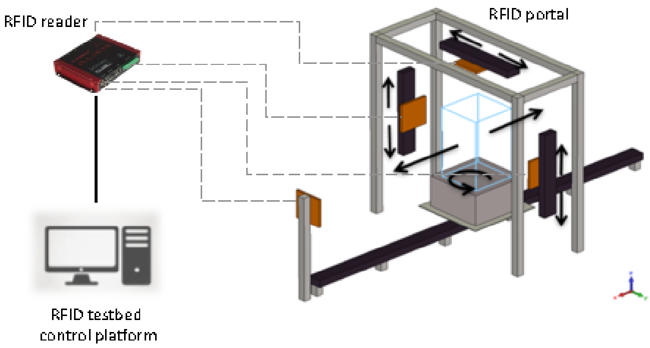


Figure 1. RFID testbed

This system allows both interactive testing as well as long running experiments of RFID reading protocols. The software platform was designed to allow fine control over all dynamic aspects influencing RFID readings: movements for target and antenna, RFID reader configuration, and smart antenna configuration (diversity and power control). Given this flexibility, this platform should be able to reproduce most of the situations found in real applications. In particular, it can be used to design custom reading set up optimized for various RFID portal applications [3].

6.3. Context-aware dynamic Smart Home Platform

Participants: Andrey Boytsov, Aurélien Richez, Yoann Maurel, Frédéric Weis [contact].

Tacoma group is focussed on the conception and implementation of innovative services for the Smart Home. The range of considered services is broad : from "optimizing the energy consumption" to "helping users to find their way in a building". To provide such services, automation based on pre-set scenarios is ineffective: human behavior is hardly predictable and application should be able to adapt their behavior at runtime depending on the context. We focused on recognizing user's activities to adapt applications behaviours.

Building efficient and accurate context awareness was and is still a great challenge but we proved, through the use of dedicated algorithms and a layered architecture that it is achievable when the targeted Home is known - due to the specific and non automated calibration process we used. Among all the available theories, we decided to use the Belief Function Theory (BFT) [8] [9] as it allows to express uncertainty

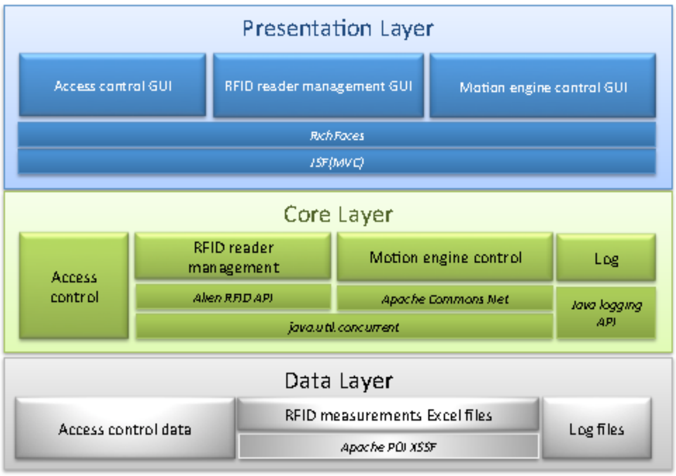


Figure 2. Software architecture of the RFID testbed

and imprecision. Although these results are very promising, great challenges still lied in (i) the support of the dynamic reconfiguration to face evolving hardware or software conditions and (ii) the deployment and the configuration of the layered architecture and sensors to allow the use of our approach in unknown environments.

One of our goals is to build a pervasive platform with constrained performance and cost [7]. The cost is particularly critical for sensors and actuators: we choose to limit our scope to inexpensive and non-invasive sensors *i.e.* no video camera. This past months, Tacoma has been working on the conception and implementation of a Smart Home Platform based on earlier prototypes inherited from ACES team. The prototypes were implemented as an hard-to-maintain monolithic code. The code also suffered from a lot of redundancy. More importantly the platform hardly supported dynamism and provided no support for reconfiguration and adaptation at runtime. With this in mind, during the re-writing of the platform the emphasis has been placed on the following aspects:

- supporting the dynamic discovery of heterogeneous sensors;
- enabling the dynamic deployment of applications at runtime ;
- enabling context-awareness by providing contextual information to these applications;
- enhancing the separation of concerns and code-reuse.

Our goal is to design and build a platform that is:

- **evolutive**: the Home environment is ever-changing and thus it is important to allow users to add new sensors or new services dynamically at runtime. It is also mandatory to recalibrate the sensors to face the change in the Home. This is mainly why we based our platform on OSGi;
- **maintainable and administrable**: we raised the maintainability by using a modular approach using C-modules or iPOJO components; the platform is itself modular to achieve a good separation of concerns (e.g., communication, module loading, discovery). We also built in-production monitoring interfaces that provides information on the belief functions that are used, the fusion process and the sensors values;
- **easy to configure**: alleviating the complexity of the platform configuration and maintenance is a prerequisite for the adoption of Smart-Home environments by consumers. Currently the BFT theories requires a huge calibration process. We focussed our efforts on the semi-automated building of mass functions, required by the theory, that have to be provided by each sensor.

6.3.1. Towards dynamism using OSGI

The development of our initial platform in C proved to be costly and hard to maintain. The dynamism is hard to achieved with a low-level language and requires an heavy development process. This led the team to investigate the use of OSGi as a based for our execution platform. OSGi is the specification of an execution framework developed on top of Java. It relies on the Java's dynamic features (dynamic and on demand class loading through class loaders) to provide a coarse-grained level of modularity. This choice was reinforced by our collaboration with the Adele team (LIG Laboratory in Grenoble). This team is using OSGi as a core for building Smart Home applications. Using OSGi would ease collaboration and code sharing.

One main concerns regarding the use of Java was the limited performances of the targeted hardware (raspberry pi). The Belief Function Theory (BFT) requires heavy computations and the embedded CPU could have been the bottleneck. Moreover, the JVM supported by the raspberry pi is limited compared to standard JVM. As a preliminary study, we choose to implement the core of the BFT library in Java and to compare the performances with the C implementation. Unexpectedly the Java implementation performed better than the C implementation in most of the case. This can be explained by three factors. First, the BFT theory is tedious to implement in low-level language. The C-implementation could probably be optimized but this will lower the readability of the source code and impact the maintainability. Second and conversely, using Java raised the code readability and allowed us to performed some optimization. Third, the JIT (Just In Time) compiler provided by the VM have been improved these past years and the optimization performed by the VM are sufficient to bring on par performances with the C implementation. As, the performances of the C platform were largely sufficient, this preliminary phase validated our decision to switch to OSGi.

6.3.2. Automated configuration of sensors

A previous defended in the group in december 2013 has shown promising results applying the BFT theory to the Smart Home Domain. It is currently possible to collect sensor values and extract belief functions from them. The platform can then extract a context from the belief functions and offer services to the user depending on what is happening. For instance, the user may be notified of an open window when he leaves the house.

The transition between a raw sensor value and a belief function is made through the use of a belief model which maps a sensor value to a belief function. The belief model is provided to the platform by us and a component is in charge of transforming a sensor value in a belief function. The fine tuning of a model can be a tedious task. It must be done by a specialist who understands the belief function theory and knows the behavior of the sensors. The model is often built iteratively by experimenting. This may take several hours or days.

Ideally, the calibration of the model should be as automatic as possible (few interaction with the user during calibration). The person setting up the sensors should not have to understand the belief function theory. The group is currently studying the possible use of clustering and classifications algorithm in order to ease the calibration of sensors. Yoann Maurel and Frédéric Weis supervised a project with a group of ENS student on this subject. The goal is to generate our belief model from a training set of sensor data. We mainly focus on two algorithms: k-nearest neighbors (KNN) and overlapping k-mean (OKM). A first experimentation with KNN and motion sensors showed that this algorithm is promising. We used a training data set to compute the presence belief model. We acquired a first set of data with someone present in the experimentation room and a second data set with nobody in the room, which gives us a labelled data set.

6.4. Towards Metamorphic Housing: the on-demand room

Participant: Michele Dominici [contact].

This research activity is supported by Fondation Rennes 1 through the chair "Smart Home and Innovation", since January 2014. During the first year, we focused on identifying the needs of the industrial partners and public authorities that fund the chair.

This activity is centered on the concept of metamorphic housing (see section 4.2). During this year, we introduced a solution of metamorphic housing addressing the goals of saving space and energy in an apartment building, while preserving residents' comfort: the on-demand room. It consists in a space that is physically shared by a small group of apartments, but is assigned for the sole use of one or few particular ones at the time, as illustrated in Figure 3 . The room is designed so as to make occupants feel they did not leave their apartment at all. They seamlessly move from their dwelling to the on-demand room and conversely, without noticing the difference, as the room adapts to their preferences.

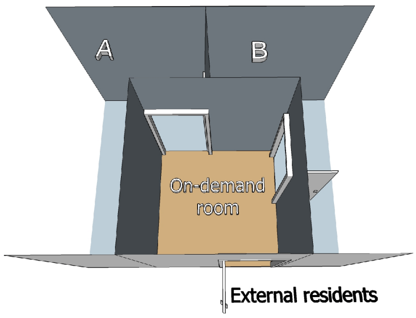


Figure 3. Floor plan for a metamorphic house

The underlying research problems are numerous. Dynamically "plugging" the room into a different apartment requires replacing the owner of the room's equipment, including appliances, heating, ventilation and air conditioning systems (HVAC), sensors, etc. The rights to control them and receive information from them must be dynamically reallocated. This must be done in a transparent fashion, so that off-the-shelf devices and appliances can be used.

In some cases, devices require dynamic reprogramming, like HVAC systems, because they must adapt to occupants' preferences and settings (e.g., ambient temperature set point).

Another research problem is the automatic learning of a schedule for the on-demand room. Regularities in users' requests for the room, duration of their occupation and privacy level can be discovered and learned. In this way, users do not have to manually book the room and usage conflicts can be prevented. We started investigating these research problems with an interdisciplinary approach and in collaboration with companies and public authorities [6]. We also started working on a prototype of the on-demand room solution, which will be presented as an immersive interactive virtual-reality application, leveraging the Immersia platform http://www.irisa.fr/immersia/.

TAO Project-Team

6. New Results

6.1. Highlights of the Year

- The European commission has chosen Crystal-Supergrids (http://www.artelys.com/news/120/90/Energy-The-European-Commission-Chooses-Artelys-Crystal) for energy modeling and planning in Europe. Crystal-Supergrids is based on the Post project, an ADEME project between Artelys and Inria-TAO.
- The HiggsML challenge was the all-time most popular challenge organized by Kaggle. Cécile Germain-Renaud, Balázs Kégl and Marc Schoenauer were part of the organizing committee.
- Creation of the Center for Data Science, an interdisciplinary institute of the Université Paris-Saclay. Co-chaired by Balázs Kégl, with more than 250 permanent researchers in 35 laboratories, the CDS organizes continued cross-fertilization of machine learning and domain sciences.
- Best Paper Award at PPSN .

BEST PAPERS AWARDS :

[36] **13th International Conference on Parallel Problem Solving from Nature**. I. LOSHCHILOV, M. SCHOENAUER, M. SEBAG, N. HANSEN.

6.2. Optimal Decision Making under Uncertainty

Participants: Olivier Teytaud [correspondent], Jean-Joseph Christophe, Jérémie Decock, Nicolas Galichet, Marc Schoenauer, Michèle Sebag, Weijia Wang.

The UCT-SIG works on sequential optimization problems, where a decision has to be made at each time step along a finite time horizon, and the underlying problem involves uncertainties along an either adversarial or stochastic setting.

After several years of success in the domain of GO, the most prominent application domain here is now energy management, at various time scales, and more generally planning. Furthermore, the work in this SIG has also lead to advances in continuous optimization at large, that somehow overlap with the work in the OPT-SIG (see 6.3).

The main advances done this year include:

Bandit-based Algorithms Active learning for the identification of biological dynamical systems has been tackled using Multi-Armed Bandit algorithms [35]. Weijia Wang's PhD [5] somehow summarizes the work done in TAO regarding Multi-objective Reinforcement Learning with MCTS algorithm. Differential Evolution was applied as an alternative to solve non-stationary Bandit problems [45].

Continuous optimization: parallelism, real-world, high-dimension and cutting-plane methods Our work in continuous optimization extends testbeds as follows: (i) including higher dimension (many testbeds in evolutionary algorithms consider dimension ≤ 40 or ≤ 100) (ii) taking into account computation time and not only the number of function evaluations (this makes a big difference in high dimension) (iii) including real world objective functions (iv) including parallelism, in particular, parallel convergence rates for differential evolution and particle swarm optimization [21]. We have a parallel version of cutting plane methods, which use more than black-box evaluations of the objective functions - we keep in mind that some of our black-box methods, on the other hand, also do not need convexity or the existence of a gradient.

Noisy optimization We have been working on noisy optimization in discrete and continuous domains. In the discrete case, we have shown the impact of heavy tails, and we have shown that resampling can solve some published open problems in an anytime manner. In the continuous case, we have shown [16] that a classical evolutionary principle (namely the step-size proportional to the distance to the optimum) implies that the optimal rates can not be reached - more precisely, we can have simple regret at best $O(1/\sqrt{number\ of\ fitness\ evaluations})$ in the simple case of an additive noise, whereas some published algorithms reached $O(1/number\ of\ fitness\ evaluations)$. One of the most directly applicable of our works is bias correction when the objective function $f(x)$ has the form $f(x) = \mathbb{E}_\omega f(x, \omega)$ and is approximated by $f(x) = \frac{1}{N}\sum_{i=1}^{N} f(x, \omega_i)$ for a given finite sample $\omega_1, \cdots, \omega_N$. We have also worked on portfolios of noisy optimizers [20], [34].

Discrete-time control with constrained action spaces. While Direct Policy Search is a reliable approach for discrete time control, it is not easily applicable in the case of a constrained high-dimensional action space. In the past, we have proposed DVS (Direct Value Search) for such cases [54]. The method is satisfactory, and we have additional mathematical results; in particular we prove positive results for non-Markovian, non-convex problems, and we prove a polynomial-time decision making and, simultaneously, exact asymptotic consistency for a non-linear transition [24]. Related work [60] also proposes to directly learn the value function, in a RL context, using some trajectories known to be bad.

Games. While still lightly contributing to the game of GO with our taiwanese partners [8], we obtained significant improvements in randomized artificial intelligence algorithms by decomposing the variance of the result into (i) the random seed (ii) the other random contributions such as the random seed of the opponent and/or the random part in the game. By optimizing our probability distribution on random seeds, we get significant improvements in e.g. phantom Go. This is basically a simple tool for learning opening books [44].

Adversarial bandits. High-dimensional adversarial bandits lead to two main drawbacks: (i) computation time (ii) highly mixed nature of the obtained solution. We developped methods which focus on sparse solution. Provably consistent, these methods are faster when the Nash equilibrium is sparse, and provides highly sparse solutions[17].

6.3. Continuous Optimization

Participants: Ouassim Ait Elhara, Asma Atamna, Anne Auger, Alexandre Chotard, Nikolaus Hansen, Yann Ollivier, Marc Schoenauer, Michèle Sebag, Olivier Teytaud, Luigi Malago, Emmanuel Benazera.

Our main expertise in continuous optimization is on stochastic search algorithms. We address theory, algorithm design, and applications. The methods we investigate are adaptive techniques able to learn iteratively parameters of the distribution used to sample solutions. The Covariance Matrix Adaptation Evolution Strategy (CMA-ES) is nowadays one of the most powerful methods for derivative-free continuous optimization. We work on different variants of the CMA-ES to improve it in various contexts as described below. We are well recognized in the field and were invited to write a book chapter on the design of continuous stochastic search[50].

Online adaptation of CMA-ES hyperparameters CMA-ES uses clever mechanisms to adapt the covariance matrix and the step-size, based on the evolution path. However, these mechanisms in turn use learning parameters, that were adjusted by trial-and-error in the seminal algorithm. However, thanks to the invariance properties of the algorithm, these values have been demonstrated to be very robust. An original mechanism has been proposed to adapt these hyper-parameters online, maximizing the likelihood of the selected sample at time to adapt the hyperparameters at time t-1. The corresponding paper published at PPSN received the Best Paper Award [36].

Linear Time and Space Complexity CMA-ES for Large-Scale Optimization We have been proposing a large-scale version of CMA-ES where the covariance matrix is restriced to a linear numbers of parameters. The update for the covariance matrix has been derived using the Information

Geometric Optimization (IGO) framework and cumulation concepts borrowed from the original CMA have been additionally included [14]. This work is part of a joint project between the TAO team and Shinshu university in Japan funded by the Japanese governement. In this context, Luigi Malago is visiting the team working on extending the proposed algorithm to a richer model.

Evaluation of Black-Box Optimizers We have been focusing on appraising the performance of step-size adaptation mechanisms for stochastic adaptive algorithms. We have shown that a too restrictive choice of test functions for the design of a method leads to misleading conclusions and proposed a thorough framework for evaluating step-size mechanism [29]. We have been pursuing our effort for *thorough and rigorous* benchmarking of black-box algorithms by organizing two more Black-Box-Optimization Benchmarking workshops that will take place at CEC 2015 and GECCO 2015. Those workshops are based on the platform COCO that we develop in the context of the ANR NumBBO project.

Theoretical Analysis of Stochastic Adaptive Algorithms We have analyzed the CSA-ES algorithm using resampling for **constrained optimization** on a linear function with a linear contraint. We have studied the behavior of the algorithm and proven success of failure of the algorithm depending on internal parameters of the algorithm [22]. We have extended a previous work on a linear function from the use of standard normal distribution to more general ones [23]. The published paper has been invited for an extension in an ECJ special issue. The extended paper had been submitted in december 2014. We have been providing a general methodology to prove the linear convergence of Comparison-based Step-size Adaptive Randomized Search on scaling-invariant functions by analyzing the stability of underlying Markov chains [57].

CMA-ES Library Besides our continuous work on implementations of CMA-ES (see e.g. github, PyPI), we have created a new library in C++11 (libcmaes). As part of the ANR SIMINOLE project, the library has been coupled with ROOT, the data analysis framework used at CERN, and generally in physics.

6.4. Applications to E-science

Participants: Cécile Germain-Renaud [correspondent], Marco Bressan, Philippe Caillou, Dawei Feng, Cyril Furtlehner, Blaise Hanczar, Karima Rafes, Balázs Kégl, Michèle Sebag.

The E-S-SIG explores the issues related to applications to E-Science, starting with modeling and optimizing very large scale computational grids, in particular in the context of Physics, to social sciences modelling with multi-agent systems.

The Higgs boson Machine Learning challenge The HiggsML challenge [0] has been set up to promote collaboration between high-energy physicists and computer scientists. The challenge, hosted by Kaggle, has drawn a remarkably large audience (with 1700+ teams it is one of the all-time most popular Kaggle challenges) and large coverage both in the social networks and in the media.

The goal of the challenge is to improve the procedure that classifies events produced by the decay of the Higgs boson versus events produced by other (background) processes, based on a training set of 250,000 examples. The challenge is a premier: it is the first time that a CERN experiment (ATLAS) made public such a large set of the official event and detector simulations. It also features a unique formal objective representing an approximation of the median significance (AMS) of a discovery (counting) test, which generates interesting algorithmic/theoretical questions beyond the usual challenges of finding and tuning the best classification algorithm [55].

A follow-up, the HEPML workshop was organized at NIPS14 [0], reporting on the results and the winning algorithms. The dataset and a software toolkit are available from the CERN Data Portal [0]

[0] https://www.kaggle.com/c/higgs-boson
[0] http://nips.cc/Conferences/2014/Program/event.php?ID=4292
[0] http://opendata.cern.ch

The Center for Data Science is a Lidex of the Université Paris-Saclay (UPSay), headed by Balazs Kégl and Arnak Dalalyan, gathering over 52 research teams and 34 labs with the goal of designing and applying automated methods to analyze massive and complex scientific datasets in order to extract useful information. Data science projects require expertise from a vast spectrum of disciplines (statistics, signal processing, machine learning, data mining, data visualization, high performance computing), besides the mastery of the scientific domain where the data originate from.

The goal of CDS is to establish an institutionalized agora in which scientists can find each other, exchange ideas, initiate and nurture interdisciplinary projects, and share their experience on past data science projects. To foster synergy between data analysts and data producers CDS organizes actions to provide initial resources for helping collaborations to get off the ground, to mitigate the non-negligible risk taken by researchers venturing into interdisciplinary data science projects, and to encourage the use of unconventional forms of information transmission and dissemination essential in this communication-intensive research area. The CDS fits perfectly in the recent surge of similar initiatives, both at the international and at the national level, and it has the potential to make the University Paris-Saclay one of the international fore-runners of data science [0].

Fault management As Lamport formulated decades ago, fault management in distributed systems exemplifies the unreachability of exact prior knowledge. Real-world large scale system add a supplementary complexity, which is non-stationarity.

- [12] models the system state and its ruptures (non-stationarity) through the flow of jobs as a stream (scalability), with a traceability goal (interpretability). These new streaming approaches involve self-calibration of the model based on scale invariance.
- D. Feng's PhD thesis [3] formulates the problem of probe selection for fault prediction based on end-to-end probing as a Collaborative Prediction (CP) problem, based on the reasonable assumption of an underlying factorial model. [26] extends the matrix completion/compressd sensing setup to a sequential (tensor) context. We propose and evaluate a new algorithme, *Sequential Matrix Factorization* (SMF) that combines matrix completion with a self-calibrating exploration/exploitation balancing heuristic. Its active learning version (SMFA) exhibits superior performance over state-of-the-art methods.

Distributed system observation The work on distributed system automated analysis and description[7] has been persued thru the continued development of the GAMA multi-agent framework https://code.google.com/p/gama-platform/wiki/GAMA. Philipps Caillou is associated to the new young researcher ANR ACTEUR, coordinated by Patrick Taillandier (IDEES, Rouen university), which will give an additional strucuture for further collaborations.

Identifying leaders in Social Networks The Modyrum contract with the SME Augure (funding Marco Bressan's Post-doc) aims at providing criteria to identify the trend leaders from blogs, tweets and other web-site posts. The same methods is being applied to fashion leaders in business as well as to opinion leaders in politics.

6.5. Designing criteria

Participants: Jamal Atif, Aurélien Decelle, Cyril Furtlehner, Yoann Isaac, Alexandre Quemy, Yann Ollivier, Marc Schoenauer, Michèle Sebag.

This SIG, rooted on the claim that *What matters is the criterion*, aims at defining new learning or optimization objectives reflecting fundamental properties of the model, the problem or the expert prior knowledge.

A statistical physics perspective With motivating applications in large scale inference problems like traffic congestions we are pursuing our quest of practical solutions to inverse problems like in [39] where a method is proposed to invert a Gaussian Markov random field with topological and spectral constraints well suited to subsequent use of belief propagation as inference algorithm

[0] http://www.datascience-paris-saclay.fr/en

(see https://who.rocq.inria.fr/Jean-Marc.Lasgouttes/star-ips for the implementation). A more specific model for traffic inference has also been developped in [11]. A method adapted to the generalized belief propagation framework, aiming at adressing directly and systematically the loop corrections without loss of scalability is about to be completed.

Multi-objective ATC The new Bayesian approach of Air Traffic Control belongs to this SIG, but was described in the Section 4.2 . Main publications are Gaétan Marceau's PhD [4] and the corresponding PPSN paper [38], [59].

Programming by Feedback Riad Akrour's PhD work on Preference Based Learning [1] culminated with the addition of a model for the user's competence in the interactive learning loop. In the resulting original paradigm, the user is sequentially proposed a series of behaviors and is only asked "Hot-or-cold" questions. The *Programming by Feedback* paradigm [15] will hopefully initiate a general way to allow non-digitally-proficient users to nevertheless control the behavior of software-based agents in their environment.

Multi-objective AI Planning This activity had almost stopped since the end of the DESCARWIN ANR project. However, a productive intership resulted in some new benchmarks in the ZenoTravel domain together with an exact solver ensuring the knowledge of the true Pareto front [41], [40].

Algorithm Selection Algorithm Selection can be viewed as a Collaborative Filtering problem, in which a problem "likes" an algorithm that is able to solve it. Initiated during Mustafa Misir's ERCIM postdoc in 2013, this idea has also been applied for Process Management [43], and is the basis of François Gonards's PhD funded by IRT SystemX in the context of aeronautics and car industry.

Outlier rejection in classification An original approach based on One-Class SVM has been proposed during Blaise Hanczar's on year *delegation* at TAO [28].

Learning sparse representations by auto-encoders Auto-encoders (AE) are a widely used tool for unsupervised learning, which consists of a neural network trained to reconstruct its own input via smaller-dimensional layers. The usual training criterion is the reconstruction error, however, the usual justification for AE is to learn a more compact data representation. In [62] we formalize this latter criterion using Minimum Description Length (MDL) and establish a comparison with the traditional reconstruction criterion. The MDL criterion has an interpretation as a denoising reconstruction and fully determines an optimal noise level, contrary to the literature on denoising AEs. More surprisingly, AE (aka Auto-associators) can also be used to learn sparse representations in the context of supervised learning [51].

TASC Project-Team

6. New Results

6.1. Highlights of the Year

In the context of the MiniZinc Challenge and in concurrency with 16 other solvers, CHOCO has won three bronze medals in three out of four categories: free search, parallel search and Open class.

6.2. CHOCO

Participants: Jean-Guillaume Fages, Narendra Jussien, Xavier Lorca, Charles Prud'Homme.

- For second consecutive year, CHOCO has participated at the MiniZinc Challenge, an annual competition of constraint programming solvers. In concurrency with 16 other solvers, CHOCO has won three bronze medals in three out of four categories (Free search, Parallel search and Open class). Five versions have been released all year long, the last one (v3.3.0, Dec. 17th) has the particularity to be promoted on Maven Central Repository. The major modifications were related to a simplification of the API but also improvement of the overall solver.
- Within the context of the PhD thesis of Charles Prud'homme [15], a domain specific language that allows prototyping propagation engines was integrated within CHOCO, A paper appears at Constraints.
- Within the context of the PhD thesis of Charles Prud'homme [15], a generic strategy based on explanations for large neighborhood search was designed and integrated within CHOCO. A corresponding paper appears at Constraints [23].
- Within the context of the PhD thesis of Jean-Guillaume Fages, a documented package for graph variables was designed and integrated within CHOCO.

6.3. IBEX Solver

Participants: Gilles Chabert, Alexandre Goldsztejn, Bertrand Neveu, Gilles Trombettoni.

In 2014 the development on IBEX has focused on the following points:

- Rejection test based on first-order conditions (see First Order Rejection Tests For Multiple-Objective Optimization, A. Goldsztejn et al. [42]).
- Q-intersection (see Q-intersection Algorithms for Constraint-Based Robust Parameter Estimation, C. Carbonnel et al., AAAI 2014)

6.4. Packing curved objects

Participants: Nicolas Beldiceanu, Gilles Chabert, Ignacio Salas Donoso.

The development of algorithms to pack curved objects has continued. The filtering algorithm developed in 2013 for generic objects shapes has been published in the CP 2014 conference. Based on this result, we have started the design of a generic (nonlinear) packing solver in 2014. The strategy for packing is directly inspired from a successful approach recently proposed by our project partners (see On solving mixed shapes packing problems by continuous, T. Martinez et al., first BRICS countries congress on Computational Intelligence). It makes use of a stochastic optimization algorithm (CMA-ES) with a fitness function that gives a violation cost and equals zero when objects are all packed. We have generalized their approach by replacing the ad-hoc formulas (for measuring the overlapping between two objects) with an automatic calculation based on our filtering algorithm. The solver is done and the experiments have started.

6.5. Robustness and scheduling

Participants: Nicolas Beldiceanu, Mats Carlsson, Alban Derrien, Arnaud Letort, Thierry Petit, Stéphane Zampelli.

- *Robustness in the Context of the Cumulative Constraint*: This research [33] investigates cumulative scheduling in uncertain environments, using constraint programming. We get a new declarative characterization of robustness, which preserves solution quality which allow adding constraints to the main problem. In this context we adapt the 2013 sweep based algorithm in order to scale and handle several thousand of activities. We highlight the significance of our framework on a crane assignment problem with business constraints.
- *Characterization of Relevant Intervals in the Context of Energetic Reasoning*: Energetic Reasoning (ER) is a powerful filtering algorithm for the Cumulative constraint. Unfortunately, ER is generally too costly to be used in practice. One reason of its bad behavior is that many intervals are considered as relevant, although most of them should be ignored. In the literature, heuristic approaches have been developed in order to reduce the number of intervals to consider, leading to a loss of filtering. We provide a sharp characterization that allows to reduce the number of intervals by a factor seven without any loss of filtering [38].
- *Fix Point over a Conjunction of Scheduling Constraints*: This research introduces a family of synchronized sweep-based filtering algorithms for handling scheduling problems involving resource and precedence constraints. The key idea is to filter all constraints of a scheduling problem in a synchronized way in order to scale better. In addition to normal filtering mode, the algorithms can run in greedy mode, in which case they perform a greedy assignment of start and end times. The filtering mode achieves a significant speed-up over the decomposition into independent cumulative and precedence constraints, while the greedy mode can handle up to 1 million tasks with 64 resource constraints and 2 million precedences. These algorithms were implemented in both CHOCO and SICStus [21].

6.6. Global constraints

Participants: Nicolas Beldiceanu, Jean-Guillaume Fages, Xavier Lorca, Thierry Petit.

- Scalability becomes more and more critical to decision support technologies. In order to address this issue in Constraint Programming, we introduce the family of self-decomposable constraints. These constraints can be satisfied by applying their own filtering algorithms on variable subsets only. We introduce a generic framework which dynamically decompose propagation, by filtering over variable subsets. Our experiments over the cumulative constraint illustrate the practical relevance of self-decomposition [34].
- Consider a constraint on a sequence of variables functionally determining a result variable that is unchanged under reversal of the sequence. Most such constraints have a compact encoding via an automaton augmented with accumulators, but it is unknown how to maintain domain consistency efficiently for most of them. Using such an automaton for such a constraint, we derive an implied constraint between the result variables for a sequence, a prefix thereof, and the corresponding suffix. We show the usefulness of this implied constraint in constraint solving, both by local search and by propagation-based systematic search [25].
- Constraints over finite sequences of variables are ubiquitous in sequencing and timetabling. This led to general modelling techniques and generic propagators, often based on deterministic finite automata (DFA) and their extensions. We consider counter-DFAs (cDFA), which provide concise models for regular counting constraints, that is constraints over the number of times a regular-language pattern occurs in a sequence. We show how to enforce domain consistency in polynomial time for at-most and at-least regular counting constraints based on the frequent case of a cDFA with only accepting states and a single counter that can be increased by transitions. We also show that the satisfaction of exact regular counting constraints is NP-hard and that an incomplete propagator for

exact regular counting constraints is faster and provides more pruning than existing propagators. Finally, by avoiding the unrolling of the cDFA used by cost regular, the space complexity is reduced[26].

6.7. Optimization

Participants: Salvador Abreu, Alejandro Reyes Amaro, Yves Caniou, Philippe Codognet, Daniel Diaz, Jean-Guillaume Fages, Xavier Lorca, Éric Monfroy, Florian Richoux, Louis-Martin Rousseau.

- The traveling salesman problem (TSP) is a challenging optimization problem for CP and OR that has many industrial applications. Its generalization to the degree constrained minimum spanning tree problem (DCMSTP) is being intensively studied by the OR community. In particular, classical solution techniques for the TSP are being progressively generalized to the DCMSTP. Recent work on cost-based relaxations has improved CP models for the TSP. However, CP search strategies have not yet been widely investigated for these problems. The contributions of this research are twofold. We first introduce a natural generalization of the weighted cycle constraint (WCC) to the DCMSTP. We then provide an extensive empirical evaluation of various search strategies. In particular, we show that significant improvement can be achieved via our graph interpretation of the state-of-the-art Last Conflict heuristic. The work was published in the Constraints journal, see the salesman and the tree: the importance of search in CP.
- In the context of nature inspired metaheuristics and its combination with CP, some new work were conducted in the field of ant colony to solve the software project scheduling problem [19], and in the field of the Manufacturing Cell Design Problem [29].
- We implement new algorithmic methods for constraint problems on massively parallel machines. In [18], we propose an extensive study of homogeneous multi-walk parallel scheme for meta-heuristics both with and without communication. The next step will be to look at heterogeneous portfolio approaches where different solvers are looking in parallel for a solution to a given problem.

6.8. Modelling

Participants: Broderick Crawford, Frédéric Lardeux, Éric Monfroy, Ricardo Soto.

- In the framework of conversion of CST set constraints to SAT instances, a filtering engine has been studied and implemented in order to reduce the size of the generated SAT instances.
- From the one hand, CSP is very expressive. On the other hand, SAT solvers can solve huge instances (millions of variables and clauses). We thus worked on the conversion of CSP set constraints into SAT instances [35]. We then focused on the Social Golfer Problem, in order to easily integrate usual improvements (such as symmetry breaking) using our framework [40].

6.9. AI for real time strategy games

Participants: Santiago Ontanon, Florian Richoux, Alberto Uriarte.

We continue to develop an artificial intelligence, AIUR, to play the real time strategy (RTS) game $StarCraft^{tm}$, using both machine learning and constraint-based techniques. AIUR finished 4^{th} over 18 finalists to the $StarCraft^{tm}$ AI competition organized at the conference AIIDE 2014, and 4^{th} over 13 finalists to the competition at CIG 2014. This year, we wrote an ad-hoc CSPsolver to deal with the wall-in optimization problem [36] for StarCraft, and generalized it as a framework enable to handle any kind of CSP/COPmodels representing a RTS-related problem. This framework, named GHOST, helps the user to implement his CSP/COPmodel before solving it with the ready-to-use, already-tuned embedded solver.

TEA Project-Team

6. New Results

6.1. Highlights of the Year

This year's effort has been mainly devoted to the successful creation of project-team TEA and the definition of its new research perspective on Time, Events and Architectures in CPS design.

The SAE committee on the AADL adopted our recommendations to implement a timed and synchronous behavioural annex [13], [11] for standardisation [20]. The specification and reference implementation of this revised behavioral annex will be the focus of most our attention next year.

Adnan Bouakaz published and implemented more of the original results from his PhD. work on abstract affine scheduling [14], [15].

6.2. Priority-Driven Scheduling of Static Dataflow Graphs through Time Abstraction

Participants: Adnan Bouakaz, Thierry Gautier, Jean-Pierre Talpin.

Static dataflow graph models, such as SDF [0] and CSDF [0], are widely used to design concurrent real-time streaming applications due to their inherent functional determinism and predictable performances. The state of the art usually advocates static-periodic scheduling of dataflow graphs over dynamic scheduling. Through the past decades, a considerable effort has been made to solve this problem [0]. Ensuring boundedness and liveness is the essence of the proposed algorithms in addition to optimizing some nonfunctional performance metrics (e.g. buffer minimization, throughput maximization, etc.).

Nowadays real-time streaming applications on MPSoCs are increasingly complex; and runtime systems are more needed to handle resource sharing, task priorities, etc. Therefore, recent works [000] are considering dynamic scheduling policies (e.g. earliest-deadline first scheduling, deadline monotonic scheduling, etc.) for dataflow graphs. The main motivations of these works are: (1) most existing real-time operating systems support such scheduling policies; (2) applicability of the existing schedulability theory [00]; and (3) with such dynamic approach, multiple and independent applications, each designed as a dataflow graph, can run concurrently on the same platform.

Our work [00] [14], [15] proposes a sequence-based framework in which a large class of priority-driven schedules can be uniformly expressed and analyzed. Infinite sequences are used to describe the dataflow graphs (e.g. rate sequences, execution time sequences) and both concrete and abstract schedules (e.g. activation clocks, priority sequences, activation relations, etc.). The framework can be then easily adapted for specific needs (e.g.

[0] *Synchronous data-flow*. E. A. Lee and D. G. Messerschmitt. Proceedings of the IEEE, 1987.

[0] *Cycle-static data-flow*. Blisen, G. and Engels, M. and Lauwereins, R. and Peperstraete, Transactions on Signal Processing, v.2. 1996.

[0] *Software synthesis from dataflow graphs*. Battacharyya, S. and Lee, E. and Murthy, P. Kluwer Academic Publishers, 1996.

[0] *Affine Data-Flow Graphs for the Synthesis of Hard Real-Time Applications*. International Conference on Application of Concurrency to System Design. IEEE Press, 2012

[0] *Temporal analysis flow based on an enabling rate characterization for multi-rate applications executed on MPSoCs with non-starvation-free schedulers*. Hausmans, J., et al. International Workshop on Software and Compilers for Embedded Systems, 2014.

[0] *Hard-real-time scheduling of data-dependent tasks in embedded streaming applications.* Bamakhrama, M. and Stefanov, T. Embedded Systems Conference. ACM, 2011

[0] *Real time scheduling theory: a historical perspective*. Sha, L. et al. Real-Time Systems Conference. IEEE, 2004

[0] *A survey of hard real-time scheduling for multiprocessor systems.* Davis, R. and Burns, A. ACM Computing Surveys, v. 4, 2011

[0] *Buffer Minimization in Earliest-First Scheduling of Dataflow Graphs*. A. Bouakaz, J-P. Talpin. ACM conference on languages, compilers and tools for embedded systems. ACM Press, 2013.

[0] *Design of Safety-Critical Java Level 1 Applications Using Affine Abstract Clocks*. A. Bouakaz, J-P. Talpin. International Workshop on Software and Compilers for Embedded Systems, 2013.

affine scheduling). Our schedule construction approach is based on two steps. The first step consists in computing an abstract schedule which consists of a set of priority sequences, processor allocation sequences, and activation relations. An activation relation between two actors describes the relative order of their activations, and hence allows us to compute safe sizes of channels between them using worst-case overflow/underflow scenarios. This step must satisfy some correctness constraints such as consistency and exclusion of overflow and underflow exceptions. Once the best abstract schedule (w.r.t. to a performance metric) is computed, the schedule is refined by computing the actual periods and phases that ensure schedulability on the target architecture.

6.3. Formal Verification of a Synchronous Data-flow Compiler: from Signal to C

Participants: Van-Chan Ngo, Jean-Pierre Talpin, Thierry Gautier, Paul Le Guernic, Loïc Besnard.

Translation validation [0][0] is a technique that attempts to verify that program transformations preserve the program semantics. It is obvious to prove globally that the source program and its final compiled program have the same semantics. However, we believe that a better approach is to separate concerns and prove each analysis and transformation stage separately with respect to ad-hoc data-structures to carry the semantic information relevant to that phase.

In the case of the Signal compiler [1], [7][12], the preservation of the semantics can be decomposed into the preservation of clock semantics at the *clock calculation* phase and that of data dependencies at the *static scheduling* phase, and, finally, value-equivalence of variables at the *code generation* phase.

Translation Validation for Clock Transformations in a Synchronous Compiler. In this work, the clock semantics of the source and transformed programs are formally represented as *clock models*. A clock model is a first-order logic formula that characterizes the presence/absence status of all signals in a Signal program at a given instant. Given two clock models, a *clock refinement* between them is defined which expresses the semantic preservation of clock semantics. A method to check the existence of clock refinement is defined as a satisfiability problem which can be automatically and efficiently proved by a SMT solver.

Let Cp^{sig} and Val_{clk} be the functions which define the Signal compiler and a validator, respectively. The following function defines a formally verified compiler for the *clock calculation and Boolean abstraction* phase. We write $C \sqsubseteq_{clk} A$ to denote that there exists a refinement between A and C.

$$Cp^{sig}_{Val_{clk}}(A) = \begin{cases} C & \text{if } Cp^{sig}(A) = C \text{ and } Val_{clk}(A,C) = true \\ Error & \text{if } Cp^{sig}(A) = C \text{ and } Val_{clk}(A,C) = false \\ Error & \text{if } Cp^{sig}(A) = Error \end{cases}$$

where $Val_{clk}(A,C) =$ true if and only if $C \sqsubseteq_{clk} A$.

Precise Deadlock Detection for Polychronous Data-flow Specifications. Dependency graphs are a commonly used data structure to encode the streams of values in data-flow programs and play a central role in scheduling instructions during auto-mated code generation from such specifications. In this work [17], we propose a precise and effective method that combines a structure of dependency graph and first order logic formulas to check whether multi-clocked data-flow specifications are deadlock free before generating code from them. We represent the flow of values in the source programs by means of a dependency graph and attach first-order logic formulas to condition these dependencies. We use an SMT solver [0] to effectively reason about the implied formulas and check deadlock freedom.

[0] *Translation validation.* Pnueli A., Siegel M., and Singerman E. In Proceedings of TACAS'98, 1998.

[0] *Translation validation: From signal to c.* M. Siegel A. Pnueli and E. Singeman. In Correct Sytem Design Recent Insights and Advances, 2000.

[0] *Satisfiability modulo theories: An appetizer.* L. de Moura and N. Bjorner. In Brazilian Symposium on Formal Methods, 2009.

Evaluating SDVG translation validation: from Signal to C. This work focuses on proving that every output signal in the source program and the corresponding variable in the compiled program, the generated C program, have the same values. The computations of all signals and their compiled counterparts are represented by a shared value-graph, called *Synchronous Data-flow Value-Graph* (SDVG).

Given a SDVG, assume that we want to show that two variables have the same value. We simply need to check that they are represented by the same sub-graph, meaning that they point to the same graph node. If all output signals in the source program A and the corresponding variables in the generated C program have the same value, then we say that C refines A, denoted by $C \sqsubseteq_{val} A$.

Implementation and Experiments. At a high level, our tool *SigCert* (https://scm.gforge.inria.fr/svn/sigcert) developed in OCaml checks the correctness of the compilation of Signal compiler w.r.t clock semantics, data dependence, and value-equivalence as given in Figure 3 .

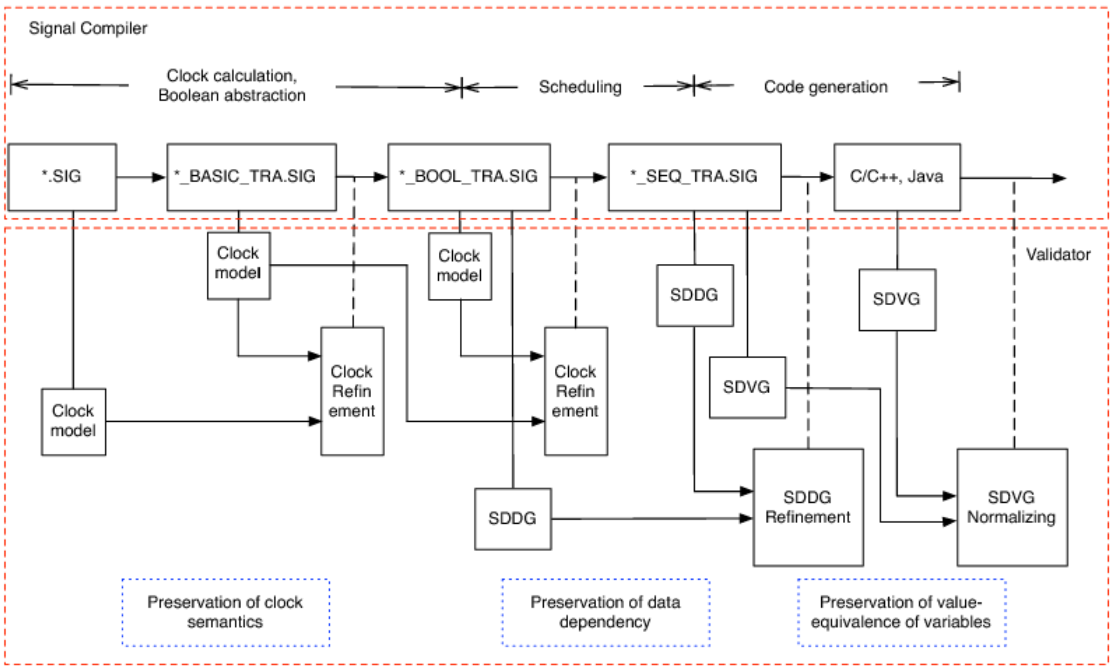


Figure 3. Our Integration within Polychrony Toolset

6.4. Ongoing integration of Polychrony with the P toolset

Participants: Christophe Junke, Loïc Besnard, Thierry Gautier, Paul Le Guernic, Jean-Pierre Talpin.

Current state of P. The FUI project P has been extended until September 2015. Partners in the project now focus on code generation aspects, leaving software architecture aspects aside. The qualifiable model-based code generator, previously known as P toolset, is now named QGen (QGen is developped mostly in Ada 2012 and Python).

Model transformation (P2S). We developped a transformation tool hereafter named P2S for expressing P system models as Signal processes. Our work is based on EMF (Eclipse Modelling Framework), taking advantage of the existing Ecore metamodels available for both P and SSME.

The P2S tool is written in Clojure, which is a dialect of Lisp running on the Java Virtual Machine. This approach allows to benefits from a terse and expressive language while remaining fully interoperable with existing Java libraries (including Eclipse plugins and especially Polychrony ones).

SSME abstraction layer.P2S uses an abstraction layer to simplify the creation of SSME elements, while taking into account EMF idioms. For example, the following expression creates a ProcessModel instance using the currently registered EMF factory:

```
(process "TestProcess"
         :in '[boolean h integer x]
         :out '[integer y]
         :body (sigdef
                (id 'y)
                (when* (id 'x) (id 'h))))
```

The newly created object can be saved as an XMI file using EMF utilities (the XMI file is 40 lines long and not shown here). This object and its children represent the following Signal process expression [0]:

```
process TestProcess =
  (? boolean h; integer x;
   ! integer y; )
  (| y := (x when h) |);
```

Transformation to P. Conversion from P to Signal relies on Clojure's multimethods. We defined a convert multimethod which dispatches on the type of its argument and possibly on additional modifiers. This mechanism allows to convert expressions differently depending on whether we want to produce a Signal declaration or an expression. For example, the following method specializer converts a P port as a signal declaration:

```
(defmethod convert [Port :declaration] [port & _]
  (ssme/signal-declarations
   (convert (.getDataType port))
     (ssme/with-comment
       [(readable-name port :declaration) :post]
       (ssme/id (p-name port)))))
```

Since the specializer contains the :declaration keyword, the previous conversion is applied only when called with that keyword given as an extra argument, as follows:

```
(convert some-port :declaration)
```

The more general specializer, which is defined below, is meant to be used inside Signal expressions and, as such, only returns a Signal identifier:

```
(defmethod convert Port [port]
   (ssme/id (p-name port)))
```

Note also that thanks to class inheritance, the above methods are sufficient to convert all kind of P ports (input/output, data/control).

The naming scheme for the resulting SSME elements is handled by the p-name multi-method and relies on XMI identifiers of the original P elements: XMI identifiers generated by QGen are string representations of positive integers. Moreover, those identifiers are guaranteed to be unique in a model. These two properties allows to generate valid Signal identifiers while ensuring traceability (e.g. signal P101 links to the unique port of the original model having 101 as a unique identifier).

Datatypes are currently converted as Signal predefineds types, which do not always match exactly the original types. Another partially implemented option consists in translating them as external types in Signal. Some types, like arrays, are converted the same way with both approaches:

[0] Even using the dedicated signalTreeAPI utility class, the same example would require many more lines of Java code.

```
(defmethod convert TArray [a]
  (reduce (fn [base dim]
            (ssme/array-type base (convert dim :signal)))
          (convert (.getBaseType a))
          (.getDimensions a)))
```

Conversion of arithmetic operations may also lead to predefined Signal operators (by default) or externally defined functions (incomplete). The current approach has been tested on QGen's test models and successfully translates 208 of the 227 models.

Partial block sequencing. The conversion from P models to Signal takes into account block dependencies as computed by QGen. Unfortunately, QGen's block sequencer produces a total order between blocks, with leads to over-constrained Signal models. We contributed to the model compiler by writing an alternative (Ada) package which provides: (i) a way to parameterize block sequencing, and (ii) partial ordering options.

Our implementation is not part of the qualified compiler, but available as a standalone (non-qualifiable) executable. However, during the development of this block sequencer, we were able to find and correct existing bugs in QGen's sequencer.

Perspectives. From a software development point of view, our current work needs to be packaged and better integrated with the build system of Polychrony. By the way, that existing build process itself could be slightly improved by using Maven configuration files instead of Eclipse manual plug-in management.

The use of a functional language on top of the Java Virtual Machine is an interesting aspect of our work. By allowing the abstraction layer, which currently works at the SSME level, to also access the existing Signal library, we could provide an API for writing and compiling Signal code using a domain-specific language expressed in Clojure (there already exist JNI bindings with the native library). This feature could help developpers hook into, or interact with, the existing Signal compiler in order to customize parts of the code generation strategies.

Regarding the P project, we still need to test code distribution strategies on industrial use-cases and determine how it can be exploited at the system-model level.

6.5. A synchronous annex for the AADL

Participants: Loïc Besnard, Thierry Gautier, Paul Le Guernic, Jean-Pierre Talpin.

The SAE committee on the AADL adopted our recommendations to implement a timed and synchronous behavioural annex for the standard [20]. The specification and reference implementation of this revised behavioral annex will be the focus of most our attention next year.

We propose a synchronous timing annex for the SAE standard AADL. Our approach consists of building a synchronous model of computation and communication that best fits the semantics and expressive capability of the AADL and its behavioral annex and yet requires little to know (syntactic) extension to it, i.e. to identify a synchronous core of the AADL (which prerequisites a formal definition of synchrony at hand) and define a formal design methodology to use the AADL in a way that supports formal analysis, verification and synthesis.

Our approach first identifies the core AADL concepts from which time events can be described. Then, is considers the behavior annex (BA) as the mean to model synchronous signals and traces through automata. Finally, we consider elements of the constraint annex to reason about abstractions of these signals and traces by clocks and relations among them. To support the formal presentation of these elements, we define a model of automata that comprises a transition system to express explicit transitions and constraints, in the form of a boolean formula on time, to implicitly constraint its behavior. The implementation of such an automaton amounts to composing its explicit transition system with that of the controller synthesised from its specified constraints.

6.6. New features of Polychrony

Participants: Loïc Besnard, Thierry Gautier, Paul Le Guernic.

Reduction of communications. We have developed, as a general functionality of the Signal toolbox, a means to reduce communications between two graphs, using assignment clocks and utility clocks.

For a given signal x, its assignment clock represents the instants at which it may be modified (otherwise than keeping its previous value $x\$$) while its utility clock in a given graph represents the instants at which it is effectively used in this graph.

Considering two graphs G_i and G_j with a signal x sent from G_i to G_j, containers are built above G_i and G_j in order to minimize the clock at which x must be communicated. On the sender side, the signal which has to be sent can be reduced to x_j with $x_j := x \texttt{ when } h$, where h is the lower bound of the assignment clock of x and the utility clock of x in G_j. On the receiver side, x is replaced in G_j by x_r with $x\_r := x_j \texttt{ default } x\_r\$$.

Note that this reduction is not always possible because it may introduce cycles between signals and clocks.

Experiments have been made on programs intended to the distribution of Quartz applications, with a gain of up to 40on some of them [18].

Polychronous automata. We have defined a new model of polychronous constrained automata that has been provided as semantic model for our proposal of an extension of the AADL behavioural annex [20]. An algebra of regular expressions is also defined to represent abstractions of constrained automata or, more specifically, their time constraints.

An experimental implementation of the semantic features of this "timing annex" will be provided through the Polychrony framework. For that purpose, representations of automata are introduced in the Signal toolbox of Polychrony. In a first step, we have decided to provide only a minimal extension of the Signal language itself. A new syntactic category of process model, which is an automaton model, has been introduced. States are described by the association of labels with subprocesses, as it is available in Signal, and transitions between states, at a given clock, are written as calls to *intrinsic* (predefined) processes. Constraints described as regular expressions on events should also be introduced using intrinsic processes.

Automata will be used in different ways related to stategies of compilation. In particular, they will serve as an alternative model for the code generation. For that purpose, polychronous programs are rewritten thanks to valuations of memorized boolean signals. The resulting partially valuated programs are the states of a control automaton.

Such techniques can be applied to implement endo-isochronous programs. Currently, code may be generated only for endochronous programs, for which clock hierarchy is a tree. Endo-isochronous programs are compositions of endochronous programs the "intersection" of which is also endochronous. For example, an automaton can be built to generate code when two signals are known to alternate.

6.7. Optimized Distribution of Synchronous Programs via a Polychronous Model

Participants: Ke Sun, Jean-Pierre Talpin, Thierry Gautier, Loïc Besnard.

We propose a distribution methodology for synchronous programs [18], applied in particular on programs written in the Quartz language [0]. The given program is first transformed into an intermediate model of guarded actions. After user-specified partitioning, the generated sub-models are transformed into equivalent Signal processes [7]. Then, the unnecessary constraints are eliminated from the processes to avoid unnecessary synchronization. Finally, within the Signal framework, the minimal frequencies of communication and computation are computed via multi-clock calculation. This operation can efficiently reduce the communication quantity and the computation load, with no change to the interface behaviors. Along this way, an optimized data-flow network over desynchronized processing locations can be constructed.

[0] *The Synchronous Programming Language Quartz.* K. Schneider, Technical Report n. 375. University of Kaiserslautern, 2009

The presented methodology has been implemented within the integrated framework Quartz/Averest + Signal/Polychrony. To illustrate and validate this methodology, a series of examples served as case studies. Each of them has been written in the Quartz language and distributed over different processing locations using the presented optimization methodology. These case studies confirm that the optimization can bring in significant communication reduction. In the sequel, the efficient utilization of distributed systems is substantially updated.

6.8. Component-based Design of Multi-rate Systems

Participants: Ke Sun, Jean-Pierre Talpin, Thierry Gautier, Loïc Besnard.

The Synchronous language Quartz is well suited for modeling mono-clocked systems. However, as based on the model of computation (MoC) synchrony, its parallelism feature excessively strengthens the synchronization. Such synchronous parallelism in particular restricts independent component design. That is, the modeling of connected components should constantly refer to each other to guarantee the achievement of desired system behavior. Hence, Quartz cannot support well the component-based system design, in particular for the distributed systems that are generally deployed over desynchronized processing locations with multi-rate clocks.

In contrast to Quartz, the polychronous language Signal is based on the MoC polychrony. As its name suggests, a polychronous program makes use of multi-rate clocks to drive its execution. One can consider that each component in the program holds its own master clock, and there is no longer a master clock for the whole program. The resulted architecture is named globally asynchronous locally synchronous (GALS) architecture.

Through integrating Quartz with Signal, a component-based methodology is proposed for designing multi-rate systems: at first, components are modeled independently to achieve local behaviors; secondly, inter-component communications are adjusted using Signal to realize intermittent synchronization. In this way, the modeling approach for mono-clocked systems evolves into a component-based modeling methodology. Such significant progress not only facilitates the component coordination, but also enhances the component reusability, in particular for modeling large scale systems.

TEMPO Team

6. New Results

6.1. Highlights of the Year

The project was created.

6.2. Approximately Timed Simulation

Participants: Vania Joloboff, Shenpeng Wang.

Existing fast simulators such as SimSoC are Loosely Timed. They evaluate the time taken by instructions executed based on an average model. Typically, the clock value is increased by a constant K every N instructions. This is sufficient to test application software with time-outs or to synchronize multicore applications, but it cannot provide a reasonable performance estimate of the embedded software.

To obtain precise parformance estimate, a common practice is to run the software on Cycle Accurate simulators, which provides a performance measure absolutely correct, but take a very long time. This is becoming a bottleneck. In fact, in many cases the software developers need some performance estimate, but do not require cycle precision. The idea of "Approximately Timed" simulation is to provide a fast simulation that can be used by software developers, and yet provide performance estimate. The goal of approximately timed simulation is to provide estimates that are within a small margin error from the real hardware, but at a simulation speed that is an order of magnitude faster than a cycle accurate one.

It is possible to maintain fast simulation, (though slower than Loosely Timed) whereas predicting reasonably accurate performance. The challenge is to come up with an abstract model of the processor that does not simulate the processor at cycle level but simulate enough to measure elapsed time with good precision. The approach is the following: a modern processor in nominal mode executes at least one instruction per clock cycle. If it does not do so, it is because there is a delay, whether a cache miss, a pipe line stall, etc. If one can simulate enough of the system so that the cause of the delays can be reproduced in the simulation and the delays evaluated, although the details of the system are not reproduced exactly, then the delays estimate may be accurate enough to provide an acceptable margin error. Moreover some of these computation can be done only once, not for each iteration of a loop.

In our work, we are considering only the processor model and we rely upon TLM interface to the interconnect for peripheral access to provides us with timing delays. We estimate the performance by using static analysis of the application control flow graph combined with a minimum of dynamic computation in order to maintain a reasonable simulation speed. We have developed such a fast Approximately Timed ISS, that does not fully simulate the hardware, yet provides good precision estimates, and does not use stastistical methods. Our approach consists in developing a higher abstraction model of the processor (than the CA models) that still executes instructions using fast SystemC/TLM code, but in parallel maintains some architecture state to measure the delays introduces by cache misses and pipe line stalls, although the pipe line is not really simulated. This work will be published in 2015 in volume 68 of the WIT Transactions on Information and Communication Technologies (ISBN 978-1-78466-054-3) [10].

6.3. Automated generation of simulator

Participants: Vania Joloboff, Shengpeng Liu.

Developing a simulator for a complete processor represents a lot of work when it is manual coding, and it is error-prone. Several efforts have been made to generate partly or entirely simulators. The dominant approach in the past years has been to use a high level description language of the processor and to generate code with the language compiler, such as LISA [18], MIMOLA [15], EXPRESSION-ADL [17]. But still, the architecture is described manually into the high level language. It is interesting to explore new architectures, but has the same issues as manual coding for simulation of commercial off-the-shelf processors such as ARM an Power architectures. Of course this approach only makes sense if the vendor has at least some semi-formal description of the architecture, which is not the case for Intel, but is the case for ARM, PowerPC and SH.

In order to automatically generate simulators from the vendor specification, we have initiated a new approach: generating the simulator from the specification of the hardware vendor as available from their web site as .pdf document. After a relatively successful initiative using ad-hoc tools, we wanted to pursue this work in a more robust and industrial context, using XML to generate an XML model of the instruction set from the vendor specification in .pdf, formalize some XML model transformations and finally generate directly the simulator code in C++. In addition, we wanted the translator to be architecture independent, not making any assumption during the translation process. However, this work is hitting difficult issues due to the fact that the vendor specification is incomplete and we have to do more manual architecture specific complements to the specification that we anticipated, which seriously weakens the project objective.

6.4. Automated Test

Participants: Mingsong Chen, Haifeng Gu, Xinqian Zhang.

Under the increasing complexity together with the time-to-market pressure, functional validation is becoming a major bottleneck of smart applications running on mobile platforms (e.g., Android, iOS).Unlike traditional software, smartphone applications are reactive and GUI (Graphical User Interface) intensive. The execution of smartphone applications heavily relies on the interactions with users. Manual GUI testing is extremely slow and unacceptably expensive in practice. However, the lack of formal models of user behaviors in the design phase hinders the automation of GUI testing (i.e., test case generation and test evaluation). While thorough test efforts are required to ensure the consistency between user behavior specifications and GUI implementations, few of existing testing approaches can automatically utilize the design phase information to test complex smartphone applications. Based on UML activity diagrams, we propose an automated GUI testing framework called ADAutomation, which supports user behavior modeling, GUI test case generation, and post-test analysis and debugging. The experiments using two industrial smartphone virtual prototypes demonstrate that our approach [16] can not only drastically reduce overall testing time, but also showed to improve the quality of designs.

6.5. SAT based bounded model checking

Participants: Mingsong Chen, Haifeng Gu, Xinqian Zhang.

SAT-based Bounded Model Checking (BMC) is promising for automated generation of directed tests. Due to the state space explosion problem, SAT-based BMC is unsuitable to handle complex properties with large SAT instances or large bounds. In this work, we propose a framework to automatically scale down the SAT falsification complexity by utilizing the decision ordering based learning from decomposed sub-properties. Our framework makes three important contributions: i) it proposes learning-oriented decomposition techniques for complex property falsification, ii) it proposes an efficient approach to accelerate the complex property falsification using the learning from decomposed sub-properties, and iii) it combines the advantages of both property decomposition and property clustering to reduce the overall test generation time. The experimental results [11] using both software and hardware benchmarks demonstrate the effectiveness of our framework.

TITANE Project-Team

6. New Results

6.1. Analysis

6.1.1. Parametric Object Detection in Large Scenes

Participant: Florent Lafarge.

Point processes are a natural extension of Markov Random Fields (MRF), designed to handle parametric objects. They have shown efficiency and competitiveness for tackling object extraction problems in vision. Simulating these stochastic models is however a difficult task. The performances of the existing samplers are limited in terms of computation time and convergence stability, especially on large scenes. We propose a new sampling procedure based on a Monte Carlo formalism [8]. Our algorithm exploits the Markovian property of point processes to perform the sampling in parallel. This procedure is embedded into a data-driven mechanism so that the points are distributed in the scene in function of spatial information extracted from the input data. The performances of the sampler are analyzed through a set of experiments on various object detection problems from large scenes, including comparisons to the existing algorithms. The sampler is also tested as optimization algorithm for MRF-based labeling problems.

6.2. Reconstruction

6.2.1. Indoor Scene Reconstruction

Participants: Sven Oesau, Florent Lafarge, Pierre Alliez.

In collaboration with EADS ASTRIUM

We contributed a method for automatic reconstruction of permanent structures of indoor scenes, such as walls, floors and ceilings, from raw point clouds acquired by laser scanners [6]. Our approach employs graph-cut to solve an inside/outside labeling of a space decomposition. To allow for an accurate reconstruction the space decomposition is aligned with permanent structures. A Hough Transform is applied for extracting the wall directions while allowing a flexible reconstruction of scenes. The graph-cut formulation takes into account data consistency through an inside/outside prediction for the cells of the space decomposition by stochastic ray casting, while favoring low geometric complexity of the model. Our algorithm produces watertight reconstructed models of multi-level buildings and complex scenes.

6.2.2. State of the Art in Surface Reconstruction from Point Clouds

Participant: Pierre Alliez.

In collaboration with Matthew Berger, Andrea Tagliasacchi, Lee Seversky, Joshua Levine, Andrei Sharf and Claudio Silva.

The area of surface reconstruction has seen substantial progress in the past two decades. The traditional problem addressed by surface reconstruction is to recover the digital representation of a physical shape that has been scanned, where the scanned data contains a wide variety of defects. While much of the earlier work has been focused on reconstructing a piece-wise smooth representation of the original shape, recent work has taken on more specialized priors to address significantly challenging data imperfections, where the reconstruction can take on different representations – not necessarily the explicit geometry. This state-of-the-art report surveys the field of surface reconstruction, providing a categorization with respect to priors, data imperfections, and reconstruction output. By considering a holistic view of surface reconstruction, this report provides a detailed characterization of the field, highlights similarities between diverse reconstruction techniques, and provides directions for future work in surface reconstruction [11].

6.2.3. Robust Shape Reconstruction and Optimal Transportation

Participants: Simon Giraudot, Pierre Alliez.

In collaboration with David Cohen-Steiner.

We describe a framework for robust shape reconstruction from raw point sets, based on optimal transportation between measures, where the input point sets are seen as distribution of masses. In addition to robustness to defect-laden point sets, hampered with noise and outliers, our approach can reconstruct smooth closed shapes as well as piecewise smooth shapes with boundaries [10].

6.3. Approximation

6.3.1. Zometool Shape Approximation

Participants: Henrik Zimmer, Florent Lafarge, Pierre Alliez.

In collaboration with Leif Kobbelt.

We contributed an algorithm that approximates 2-manifold surfaces with Zometool models while preserving their topology. Zometool is a popular hands-on mathematical modeling system used in teaching, research and for recreational model assemblies at home. This construction system relies on a single node type with a small, fixed set of directions and only nine different edge types in its basic form. While being naturally well suited for modeling symmetries, various polytopes or visualizing molecular structures, the inherent discreteness of the system poses difficult constraints on any algorithmic approach to support the modeling of freeform shapes. We contribute a set of local, topology preserving Zome mesh modification operators enabling the efficient exploration of the space of 2-manifold Zome models around a given input shape. Starting from a rough initial approximation, the operators are iteratively selected within a stochastic framework guided by an energy functional measuring the quality of the approximation. We demonstrate our approach on a number of designs and also describe parameters which are used to explore different complexities and enable coarse approximations [15].

6.3.2. CGALmesh: a Generic Framework for Delaunay Mesh Generation

Participants: Pierre Alliez, Clement Jamin.

In collaboration with Jean-Daniel Boissonnat and Mariette Yvinec.

CGALmesh is the mesh generation software package of the Computational Geometry Algorithm Library (CGAL). It generates isotropic simplicial meshes – surface triangular meshes or volume tetrahedral meshes – from input surfaces, 3D domains as well as 3D multi-domains, with or without sharp features. The underlying meshing algorithm relies on restricted Delaunay triangulations to approximate domains and surfaces, and on Delaunay refinement to ensure both approximation accuracy and mesh quality. CGALmesh provides guarantees on approximation quality as well as on the size and shape of the mesh elements. It provides four optional mesh optimization algorithms to further improve the mesh quality. A distinctive property of CGALmesh is its high flexibility with respect to the input domain representation. Such a flexibility is achieved through a careful software design, gathering into a single abstract concept, denoted by the oracle, all required interface features between the meshing engine and the input domain. We already provide oracles for domains defined by polyhedral and implicit surfaces [5].

6.3.3. Level-of-Detail Quad Meshing

Participant: David Bommes.

In collaboration with Hans-Christian Ebke and Leif Kobbelt from RWTH Aachen.

The most effective and popular tools for obtaining feature aligned quad meshes from triangular input meshes are based on cross field guided parametrization. These methods are incarnations of a conceptual three-step pipeline: (1) cross field computation, (2) field- guided surface parametrization, (3) quad mesh extraction. While in most meshing scenarios the user prescribes a desired target quad size or edge length, this information is typically taken into account from step 2 onwards only, but not in the cross field computation step. This turns into a problem in the presence of small scale geometric or topological features or noise in the input mesh: closely placed singularities are induced in the cross field, which are not properly reproducible by vertices in a quad mesh with the pre- scribed edge length, causing severe distortions or even failure of the meshing algorithm. We reformulate the construction of cross fields as well as field-guided parameterizations in a scale-aware manner which effectively suppresses densely spaced features and noise of geometric as well as topological kind. Dominant large-scale features are adequately preserved in the output by relying on the unaltered input mesh as the computational domain [4].

6.3.4. Mesh Watermarking based on a Constrained Optimization Framework

Participants: Xavier Rolland-Nevière, Pierre Alliez.

In collaboration with Technicolor and Gwenael Doerr.

A watermarking strategy for triangle surface meshes consists in modifying the vertex positions along the radial directions, in order to adjust the distribution of radial distances and thereby encode the desired payload. To guarantee that watermark embedding does not alter the center of mass, prior work formulated this task as a quadratic programming problem. We contribute a generalization of this formulation with: (i) integral reference primitives, (ii) arbitrary relocation directions to alter the vertex positions, and (iii) alternate distortion metrics to minimize the perceptual impact of the embedding process. These variants are evaluated against a range of attacks and we report both improved robustness performances, in particular for simplification attacks, and improved control over the embedding distortion [9].

6.3.5. Robust 3D Watermarking

Participants: Xavier Rolland-Nevière, Pierre Alliez.

In collaboration with Technicolor, thesis co-advised by Pierre Alliez and Gwenael Doerr.

3D models are valuable assets widely used in the industry and likely to face piracy issues. This dissertation deals with robust mesh watermarking that is used for traitor-tracing. Following a review of state-of-the-art 3D watermarking systems, the robustness of several content adaptation transforms are evaluated. An embedding domain robust against pose is investigated, with a thickness estimation based on a robust distance function to a point cloud constructed from some mesh diameters. A benchmark showcases the performance of this domain that provides a basis for robust watermarking in 3D animations. For static meshes, modulating the radial distances is an efficient approach to watermarking. It has been formulated as a quadratic programming problem minimizing the geometric distortion while embedding the payload in the radial distances. This formulation is leveraged to create a robust watermarking framework, with the integration of the spread-transform, integral reference primitives, arbitrarily selected relocation directions and alternate metrics to minimize the distortion perceived. Benchmarking results showcase the benefits of these add-ons w.r.t the fidelity vs. robustness watermarking trade-off. The watermark security is then investigated with two obfuscation mechanisms and a series of attacks that highlight the remaining limitations. A resynchronization approach is finally integrated to deal with cropping attacks. The resynchronization embeds landmarks in a configuration that conveys synchronization information that will be lost after cropping. During the decoding, this information is blindly retrieved and significant robustness improvements are achieved [2].

6.3.6. Spread Transform and Roughness-based Shaping to Improve 3D Watermarking based on Quadratic Programming

Participants: Xavier Rolland-Nevière, Pierre Alliez.

In collaboration with Technicolor and Gwenael Doerr.

Modulating the distances between the vertices and the center of mass of a triangular mesh is a popular approach to watermark 3D objects. Prior work has formulated this approach as a quadratic programming problem which minimizes the geometric distortion while embedding the watermark payload in the histogram of distances. To enhance this framework, we introduce two watermarking components, namely the spread transform and perceptual shaping based on roughness information. Benchmarking results showcase the benefits of these add-ons with respect to the fidelity-robustness trade-off [13].

TOCCATA Project-Team

6. New Results

6.1. Highlights of the Year

- The ACM Software System Award 2013 was given, during a ceremony in June 2014 in San Francisco, to the Coq proof assistant (http://awards.acm.org/software_system/). The prestigious ACM price was previously awarded to the LLVM compiler infrastructure (2012) and to the Eclipse IDE (2011). Among the 9 recipients of the 2013 award are Christine Paulin and Jean-Christophe Filliâtre, from the Toccata team.
- The *Concours Castor informatique* (http://castor-informatique.fr/) had an even larger success than in the previous years. In November 2014, more than 228,000 teenagers from over 1500 schools participated and solved the interactive tasks of the contest. Arthur Charguéraud and Sylvie Boldo, from the Toccata team, significantly contributed to the prepation of the tasks and to the organization of the contest.

6.2. Deductive Verification

- J.-C. Filliâtre, L. Gondelman, and A. Paskevich have formalized the notion of ghost code implemented in *Why3*, in a paper *The Spirit of Ghost Code* [35] presented at CAV 2014. This is an outcome of L. Gondelman's M2 internship (spring/summer 2013).
- M. Clochard published at the POPL conference a paper presenting a work done during an internship at Rice University (Houston, TX, USA) with S. Chaudhuri and A. Solar-Lezama [29]. It is a new technique for parameter synthesis under boolean and quantitative objectives. The input to the technique is a "sketch"—a program with missing numerical parameters—and a probabilistic assumption about the program's inputs. The goal is to automatically synthesize values for the parameters such that the resulting program satisfies: (1) a boolean specification, which states that the program must meet certain assertions, and (2) a quantitative specification, which assigns a real valued rating to every program and which the synthesizer is expected to optimize.
- J.-C. Filliâtre, C. Marché, and A. Paskevich, together with F. Bobot (CEA LIST), took part in the VerifyThis program verification competition, held at the 18th FM symposium in August 2012. They used *Why3* to solve three challenges (which can be found at http://fm2012.verifythis.org/challenges/), and their solutions have been published in a special issue of the journal *Software Tools for Technology Transfer* [16].
- M. Clochard developed, using *Why3*, verified implementations of several data structures, including random-access lists and ordered maps. These are derived from a common parametric implementation of self-balancing binary trees in the style of Adelson-Velskii and Landis trees (so-called AVLs). This work appeared at the VSTTE conference [30]. Its originality relies on the genericity of the specifications and the code, and the very high level of proof automation. Such a case study is aimed at illustrating the capabilities of *Why3* for designing certified libraries. Development is available from our gallery at http://toccata.lri.fr/gallery/avl.fr.html.
- S. Conchon and A. Mebsout have extended the core algorithm of the Cubicle model checker with a mechanism for inferring invariants. This new algorithm, called BRAB, is able to automatically infer invariants strong enough to prove industrial cache coherence protocols. BRAB computes over-approximations of backward reachable states that are checked to be unreachable in a finite instance of the system. These approximations (candidate invariants) are then model-checked together with the original safety properties. Completeness of the approach is ensured by a mechanism for backtracking on spurious traces introduced by too coarse approximations. Details can be found in A. Mebsout's PhD thesis [15].

- A. Charguéraud extended his tool CFML to support, in addition to the verification of the full functional correctness of a piece of code, the verification of the asymptotic complexity of the code. Even though it had been previously established that, in theory, amortized analysis can be explained as the manipulation of *time credits*, and that time credits can be encoded as resources in Separation Logic, CFML is the first practical tool to support the formal verification of amortized analyses for arbitrarily-complex pieces of code. The *time-credit* extension to CFML was put to practice to verify dynamic arrays (Julien Grangier's internship), and to verify a *chunked sequence* data structure [26], particularly challenging due to its use of Tarjan's data structural bootstrapping technique. The latter piece of work was presented in July at the workshop *Semantics of proofs and certified mathematics*, which took place at the Institut Henri Pointcaré. A paper describing the time-credit extension to CFML is under preparation.

6.3. Floating-Point and Numerical Programs

- C. Marché published in the *Science of Computer Programming* journal [22] a detailed description of an industrial research initially conducted in the context of the U3CAT project (ended in 2012) on static analysis of critical C code. The code involves floating-point computations on quaternions that should be of norm 1. Because of the round-off errors, a drift of this norm is observed over time. In this work a bound on this drift is determined and formally proved correct, using *Frama-C*, *Jessie* and *Why3*. Proofs are done using automated provers and in a few complex cases the Coq proof assistant. The published version is up to date with the recent versions of those tools, and the development is available on our gallery at http://toccata.lri.fr/gallery/quat.en.html
- S. Boldo, C. Lelay, and G. Melquiond worked on the Coquelicot library, designed to be a user-friendly Coq library about real analysis. An easier way of writing formulas and theorem statements is achieved by relying on total functions in place of dependent types for limits, derivatives, integrals, power series, and so on. To help with the proof process, the library comes with a comprehensive set of theorems and some automation. We have exercised the library on several use cases: in an exam at university entry level, for the definitions and properties of Bessel functions, and for the solution of the one-dimensional wave equation. These results are published in the journal *Mathematics in Computer Science* [19].
- S. Boldo and G. Melquiond, with J.-H. Jourdan and X. Leroy (Gallium team, Inria Paris - Rocquencourt) extended the CompCert compiler to get the first formally verified C compiler that provably preserves the semantics of floating-point programs This work, published in the *Journal of Automated Reasoning* [18], also covers the formalization of numerous algorithms of conversion between integers and floating-point numbers.
- S. Boldo, C. Lelay, and G. Melquiond, have conducted a survey on the formalization of real arithmetic and real analysis in various proof systems. This work, published in the journal *Mathematical Structures in Computer Science* [20], details the axioms, definitions, theorems, and methods of automation, available in these systems.
- É. Martin-Dorel and G. Melquiond worked on integrating the CoqInterval and CoqApprox libraries into a single package. The CoqApprox library is dedicated to computing verified Taylor models of univariate functions so as to compute approximation errors. The CoqInterval library reuses this work to automatically prove bounds on real-valued expressions. A large formalization effort took place during this work, so as to get rid of all the holes remaining in the formal proofs of CoqInterval. It was also the chance to perform a comparison between numerous decision procedures dedicated to proving nonlinear inequalities involving elementary functions. A report is available [43].
- S. Boldo, J.-C. Filliâtre, and G. Melquiond, with F. Clément and P. Weis (POMDAPI team, Inria Paris - Rocquencourt), and M. Mayero (LIPN), completed the formal proof of a numerical analysis program: the second-order centered finite-difference scheme for the one-dimensional acoustic wave. This proof was published with a focus towards numerical analysts, in the journal *Computers and Mathematics with Applications* [17].

- P. Roux formalized the influence of double rounding on the accuracy of floating-point arithmetic operators. In particular, this includes all the corner cases that were ignored from Figueroa's original pen-and-paper proof. Results appeared in the *Journal of Formalized Reasoning* [24].
- P. Roux formalized a theory of numerical analysis for bounding the round-off errors of a floating-point algorithm. This approach was applied to the formal verification of a program for checking that a matrix is semi-definite positive. The challenge here is that testing semi-definiteness involves algebraic number computations, yet it needs to be implemented using only approximate floating-point operations. A report is available [45].

6.4. Automated Reasoning

- In the context of the BWare project, aiming at using *Why3* and Alt-Ergo for discharging proof obligations generated by Atelier B, we made progress into several directions. New drivers have been designed for *Why3*, in order to use new back-end provers Zenon modulo and iProver modulo. A notion of rewrite rule was introduced into *Why3*, and a transformation for simplifying goals before sending them to back-end provers was designed. Intermediate results obtained so far in the project were presented both at the French conference AFADL [38] and at the international conference on Abstract State Machines, Alloy, B, VDM, and Z [34].

  On the side of Alt-Ergo, recent developments have been made to efficiently discharge proof obligations generated by Atelier B. This includes a new plugin architecture to facilitate experiments with different SAT engines, new heuristics to handle quantified formulas, and important modifications in its internal data structures to boost performances of core decision procedures. Benchmarks realized on more than 10,000 proof obligations generated from industrial B projects show significant improvements [33].
- C. Dross defended her PhD thesis in April 2014 [14], on the topic of automated reasoning modulo theories, and in particular the handling of quantifiers in the SMT approach. The main results of the thesis are: (1) a formal semantics of the notion of *triggers* typically used to control quantifier instantiation in SMT solvers, (2) a general setting to show how a first-order axiomatization with triggers can be proved correct, complete, and terminating, and (3) an extended DPLL(T) algorithm to integrate a first-order axiomatization with triggers as a decision procedure for the theory it defines. Significant case studies were conducted on examples coming from SPARK programs, and on the benchmarks on B set theory constructed within the BWare project.

6.5. Certification of Languages, Tools and Systems

- M. Clochard, C. Marché, and A. Paskevich developed a general setting for developing programs involving binders, using *Why3*. This approach was successfully validated on two case studies: a verified implementation of untyped lambda-calculus and a verified tableaux-based theorem prover. This work was presented at the PLPV conference in January 2014 [32].
- M. Clochard, J.-C. Filliâtre, C. Marché, and A. Paskevich developed a case study on the formalization of semantics of programming languages using *Why3*. This case study aimed at illustrating recent improvements of *Why3* regarding the support for higher-order logic features in the input logic of *Why3*, and how these are encoded into first-order logic, so that goals can be discharged by automated provers. This case study also illustrates how reasoning by induction can be done without need for interactive proofs, via the use of *lemma functions*. This works was presented at the VSTTE conference [31].
- M. Clochard and L. Gondelman developed a formalization of a simple compiler in *Why3*. It compiles a simple imperative language into assembler instructions for a stack machine. This case study was inspired by a similar example developed using Coq and interactive theorem proving. The aim is to improve significantly the degree of automation in the proofs. This is achieved by the formalization of a Hoare logic and a Weakest Precondition Calculus on assembly programs, so that the correctness of compilation is seen as a formal specification of the assembly instructions generated. This work conducted in 2014 will be presented at the JFLA conference in January 2015 [75].

- S. Dumbrava and É. Contejean, with V. Benzaken (VALS team, at LRI) proposed a *Coq* formalization of the relational data model which underlies relational database systems. More precisely, they have presented and formalized the data definition part of the model including integrity constraints. They have modelled two different query language formalisms: relational algebra and conjunctive queries. They also present logical query optimization and prove the main "database theorems": algebraic equivalences, the homomorphism theorem and conjunctive query minimization. This work has been published at ESOP 2014 [27].
- A. Charguéraud, together with the other members of the *JsCert* team have developed this year the first complete formalization of the semantics of the JavaScript programming language. This project is joint work with Philippa Gardner, Sergio Maffeis, Gareth Smith, Daniele Filaretti and Daiva Naudziuniene from Imperial College, and Alan Schmitt and Martin Bodin from Inria Rennes (see http://jscert.org). The formalization consists of a set of inductive rules translating the prose from the *ECMAScript Language Specification, version 5*, using the pretty-big-step semantics [74]. These rules can be used to formally reason about program behaviors or to establish the correctness of program transformations. In addition to the inductive rules, a reference interpreter has been proved correct. This interpreter may be used to run actual JavaScript program following the rules of the formal semantics. It has been used in particular to validate the formal semantics against official JavaScript test suites. The formalization of JavaScript has been published at POPL [28].

6.6. Miscellaneous

A. Charguéraud worked together with Umut Acar and Mike Rainey, as part of the ERC project *DeepSea*, on the developement of efficient data structures and algorithms targeting modern, shared memory multicore architectures. Two major results were obtained this year.

The first result is a sequence data structure that provides amortized constant-time access at the two ends, and logarithmic time concatenation and splitting at arbitrary positions. These operations are essential for programming efficient computation in the fork-join model. Compared with prior work, this novel sequence data structure achieves excellent constant factors, allowing it to be used as a replacement for traditional, non-splittable sequence data structures. This data structure, called *chunked sequence* due to its use of chunks (fixed-capacity arrays), has been implemented both in C++ and in OCaml. It is described in a paper published at ESA [26].

Another result by A. Charguéraud and his co-authors is the development of fast and robust parallel graph traversal algorithms, more precisely for parallel BFS and parallel DFS. The new algorithms leverage the aforementioned sequence data structure for representing the set of edges remaining to be visited. In particular, it uses the split operation for balancing the edges among the several processors involved in the computation. Compared with prior work, these new algorithms are designed to be efficient not just for particular classes of graphs, but for all input graphs. This work has not yet been published, however it is described in details in a technical report [40]. Note that these two graph algorithms, which involve nontrivial use of concurrent data structures, will be very interesting targets for formal verification.

TONUS Team

6. New Results

6.1. Highlights of the Year

We have implemented an OpenCL task graph version of our Discontinuous Galerkin solver that allows to overlap GPU computations and MPI communications. With this optimizations, we were recently able to achieve a 14 GFLOPS simulation with 8 GPUs on an electromagnetic test case. These results are included in the PhD of Thomas Strub (defence planned in March 2015) under the supervision of Philippe Helluy.

6.2. Development of semi-Lagrangian methods

Participants: Adnane Hamiaz, Michel Mehrenberger, Christophe Steiner.

6.2.1. Gyroaverage operator for a polar mesh

A direct method is proposed in [17] in the space configuration for the computation of the gyroaverage operator. It consists in integrating on the gyrocircles using interpolation operators (Hermite or cubic splines); see also [2]. Numerical comparisons with a standard method based on a Padé approximation are performed: (i) with analytical solutions; (ii) considering the 4D drift-kinetic model with one Larmor radius and (iii) on the classical linear DIII-D benchmark case. In particular, we show that in the context of a drift-kinetic simulation, the proposed method has similar computational cost as the standard method and its precision is independent of the radius. Extension to the quasi neutral equation has begun on a 4D model with one Larmor radius. We can exhibit some specific situations where the new method leads to more accurate results and we observe as predicted that the instability growth rate is stronger than for the Padé approximation. On the other hand, we have to face with more oscillations (e.g. on the boundary) of the new operator, which does not permit to replace the Padé approximation. Promising higher order Padé approximation are envisaged for the future.

6.2.2. Semi-Lagrangian simulations on curvilinear grids

Semi-Lagrangian schemes often deal with cartesian mesh; the extension to curvilinear grids is important in order to be able to deal with specific geometries and also for adapting the grid to save computational effort. This study is part of a general work on adding curvilinear capabilities for the simulation of drift kinetic and gyrokinetic equations in a semi-Lagrangian framework, and is in current development in the SeLaLib library.

Thus, in [28] semi-Lagrangian guiding center simulations are performed on sinusoidal perturbations of cartesian grids, thanks to the use of a B-spline finite element solver for the Poisson equation and the classical backward semi-Lagrangian method (BSL) for the advection. We are able to reproduce the standard Kelvin-Helmholtz instability test on such grids. When the perturbation leads to a strong distorted mesh, we observe that the solution differs if one takes standard numerical parameters that are used in the cartesian reference case. We can recover good results together with correct mass conservation, by diminishing the time step.

6.2.3. Field aligned semi-Lagrangian schemes

In [23] we introduce field aligned interpolation for Semi-Lagrangian schemes, by adapting a method developed by Hariri-Ottaviani to the semi-Lagrangian context. This approach is validated on the constant oblique advection equation and on a 4D drift kinetic model with oblique magnetic field in cylindrical geometry. The strength of this method is that one can reduce the number of points in the longitudinal direction. Extension to tokamak conguration in toroidal geometry is the next step of this study.

6.2.4. KEEN wave simulations, high order time splitting, non-uniform cubic splines

KEEN waves are non-stationary, nonlinear, self-organized asymptotic states in Vlasov plasmas (see [3]). They lie outside the precepts of linear theory or perturbative analysis, unlike electron plasma waves or ion acoustic waves. Steady state, nonlinear constructs such as BGK modes also do not apply. The range in velocity that is strongly perturbed by KEEN waves depends on the amplitude and duration of the ponderomotive force generated by two crossing laser beams, for instance, used to drive them. Smaller amplitude drives manage to devolve into multiple highly-localized vorticlets, after the drive is turned off, and may eventually succeed to coalesce into KEEN waves. Fragmentation once the drive stops, and potential eventual remerger, is a hallmark of the weakly driven cases. A fully formed (more strongly driven) KEEN wave has one dominant vortical core. But it also involves fine scale complex dynamics due to shedding and merging of smaller vortical structures with the main one. Shedding and merging of vorticlets are involved in either case, but at different rates and with different relative importance. The narrow velocity range in which one must maintain sufficient resolution in the weakly driven cases, challenges fixed velocity grid numerical schemes. What is needed is the capability of resolving locally in velocity while maintaining a coarse grid outside the highly perturbed region of phase space. We here report on a new Semi-Lagrangian Vlasov-Poisson solver based on conservative non-uniform cubic splines in velocity that tackles this problem head on. An additional feature of our approach is the use of a new high-order time-splitting scheme which allows much longer simulations per computational effort. This is needed for low amplitude runs. There, global coherent structures take a long time to set up, such as KEEN waves, if they do so at all. The new code's performance is compared to uniform grid simulations and the advantages are quantified. The birth pains associated with weakly driven KEEN waves are captured in these simulations. Canonical KEEN waves with ample drive are also treated using these advanced techniques. They will allow the efficient simulation of KEEN waves in multiple dimensions, which will be tackled next, as well as generalizations to Vlasov-Maxwell codes. These are essential for pursuing the impact of KEEN waves in high energy density plasmas and in inertial confinement fusion applications. More generally, one needs a fully-adaptive grid in- phase-space method which could handle all small vorticlet dynamics whether pealing or remerging. Such fully adaptive grids would have to be computed sparsely in order to be viable. This two-velocity grid method is a concrete and fruitful step in that direction.

6.2.5. Conservative semi-Lagrangian scheme

While developing a new semi-Lagrangian solver, the gap between a linear Landau run in 1D-1D and a 5D gyrokinetic simulation in toroidal geometry is quite huge. Intermediate test cases are welcome for testing the code. A new fully two-dimensional conservative semi-Lagrangian (CSL) method is presented in [6] and is validated on 2D polar geometries. We consider here as building block, a 2D guiding-center type equation on an annulus and apply it on two test cases. First, we revisit a 2D test case previously done with a PIC approach and detail the boundary conditions. Second, we consider a 4D drift-kinetic slab simulation. In both cases, the new method appears to be a good alternative to deal with this type of models since it improves the lack of mass conservation of the standard semi-Lagrangian (BSL) method.

6.3. Reduced Vlasov-Maxwell modeling

Participants: Philippe Helluy, Laurent Navoret, Thi Trang Nhung Pham.

We have tested several preliminary methods for reducing the complexity of the Vlasov equation. By expanding the distribution function on velocity basis we obtain a space-only hyperbolic system. This system takes advantage of interesting conservation or entropy properties. Several types of basis can be used: Fourier [14], piecewise Lagrange [20], [13], etc. The method has been implemented for 4D problems in the Selalib library. The next step would be to adapt the size of the expansion according to the nature of the flow region and to apply the method to the gyrokinetic model.

6.4. GPU Optimization of Discontinuous Galerkin solvers

Participants: Michaël Gutnic, Philippe Helluy, Michel Massaro, Thomas Strub.

We have continued to investigate implementations of numerical schemes on new hybrid computer architectures. We have for instance applied a very efficient Strang splitting algorithm for the numerical resolution of the MHD or compressible multiphase model ([16], [18], [22]). We have also highly optimized our DG solver CLAC ([20]) for electromagnetic applications. For instance, we have implemented an OpenCL task graph that allows to overlap GPU computations and MPI communications. With this optimizations, we were recently able to achieve a 14 GFLOPS simulation with 8 GPUs on an electromagnetic test case. These results are included in the PhD of Thomas Strub (defence planned in March 2015).

6.5. Numerical and theoretical study of reduced MHD problems for the JOREK code

Participant: Emmanuel Franck.

The Jorek code is a parallel finit element code (used at the CEA Cadarache and the IPP) which simulates the edge instabilities in the Tokamak solving reduced MHD models. Firstly we have written a family of full MHD models (resistive, diamagnetic and extended MHD models). Using this, we write the reduced MHD models close to the models implemented in the code which conserve the energy and are more stable ([35]). This work will probably be published as an Inria report next year. The second part of this work consists in writing a simplified version of the JOREK code which will be useful to test and validate future numerical research in the JOREK context. Actually we have written a code which solve simple elliptic equations in 3D toroidal geometry using Bezier, splines and Fourier expansion. The integration of simple wave model and reduced MHD models [33] is in progress. When these model will be implemented, we will test a new preconditioning for the JOREK code in these simple configurations.

6.6. Simulations of highly oscillatory Vlasov-type models

Participants: Emmanuel Frénod [Univ. Bretagne-Sud], Sever Hirstoaga.

We continued our exploration of a new time-stepping method based on an exponential integrator.

First, we have improved the algorithm introduced in [11] for solving a multi-scale 1d-1d Vlasov-Poisson system within a Particle-In-Cell method, in order to do accurate long time simulations. As an exponential integrator, the new scheme (see [10]) allows to use large time steps compared to the size of oscillations in the solution. More precisely, the new idea is to push each particle with its computed period. Our simulations show that using precise periods for each particle and at each macroscopic time step results in a more accurate scheme in long times.

Then, similar ideas are used for a 2d-2d multi-scale Vlasov-Poisson system (see [27]). We propose in a Particle-In-Cell framework a robust time-stepping method that works uniformly when the small parameter (the smallest scale) vanishes. We first verify our scheme in the framework of a proposed analytic solution with fast oscillations in time and we show that the scheme works for any initial condition. Then we test the method in the nonlinear case of a Vlasov-Poisson simulation. The scheme is able to use large time steps with respect to the typical size of the solution's fast oscillations. In addition, we show numerically that the method has accurate long time behaviour and that it is asymptotic preserving with respect to the limiting Guiding Center system.

TOSCA Project-Team

6. New Results

6.1. Probabilistic numerical methods, stochastic modelling and applications

Participants: Mireille Bossy, Nicolas Champagnat, Julien Claisse, Madalina Deaconu, Benoît Henry, James Inglis, Antoine Lejay, Oana Valeria Lupascu, Sylvain Maire, Sebastian Niklitschek Soto, Denis Talay, Etienne Tanré, Denis Villemonais.

6.1.1. Published works and preprints

- M. Bossy and J.-F. Jabir (University of Valparaíso) [13] have proved the well-posedness of a conditional McKean Lagrangian stochastic model, endowed with the specular boundary condition, and further the mean no-permeability condition, in a smooth bounded confinement domain $\mathcal{D}$.
- M. Bossy, N. Champagnat, S. Maire and L. Violeau worked with H. Leman (CMAP, Ecole Polytechnique) and M. Yvinec (Inria Sophia, EPI GEOMETRICA) on Monte Carlo methods for the linear and non-linear Poisson-Boltzmann equations [12]. These methods are based on walk on spheres algorithm, simulation of diffusion processes driven by their local time, and branching Brownian motion. Their code for the linear equation can deal with bio-molecules of arbitrary sizes, based on computational geometry tools from the CGAL C++ Library developed by the GEOMETRICA team. The non-linear equation is solved using branching Brownian motion.
- M. Bossy, O. Faugeras (Inria Sophia, EPI NEUROMATHCOMP), and D. Talay have clarified the well-posedness of the limit equations to the mean-field N-neuron models proposed in [42] and proven the associated propagation of chaos property. They also have completed the modeling issue in [42] by discussing the well-posedness of the stochastic differential equations which govern the behavior of the ion channels and the amount of available neurotransmitters. See [29].
- N. Champagnat and D. Villemonais obtained criterions for existence and uniqueness of quasi-stationary distributions and Q-processes for general absorbed Markov processes [31]. A quasi-stationary distribution is a stationary distribution conditionally on non-absorbtion, and the Q-process is defined as the original Markov process conditioned to never be absorbed. The criterion that they obtain ensures exponential convergence of the conditioned t-marginal of the process conditioned not to be absorbed at time t, to the quasi-stationary distribution and also the exponential ergodicity of the Q-process.
- M. Deaconu and S. Herrmann continued and completed the study of the simulation of the hitting time of some given boundary for Bessel processes. They constructed an original approximation method for hitting times of a given threshold by Bessel processes with non-integer dimension. In this work, they combine the additivity property of the laws of squared Bessel processes with their previous results on the simulation of hitting times of Bessel processes with integer dimension, based on the method of images and on the connexion with the Euclidiean norm of the Brownian motion [33].
- M. Deaconu, S. Herrmann and S. Maire introduced a new method for the simulation of the exit time and position of a δ-dimensional Brownian motion from a domain. The main interest of this method is that it avoids splitting time schemes as well as inversion of complicated series. The idea is to use the connexion between the δ-dimensional Bessel process and the δ-dimensional Brownian motion thanks to an explicit Bessel hitting time distribution associated with a particular curved boundary. This allows to build a fast and accurate numerical scheme for approximating the hitting time [34].
- M. Deaconu and O. Lupaşcu worked with L. Beznea (Bucharest, Romania) on the construction and the branching properties of the solution of the fragmentation equation and properly associate a continuous time càdlàg Markov process. The construction and the proof of the path regularity of the Markov processes are based on several newly developed potential theoretic tools.

- J. Inglis, together with O. Faugeras (Inria NeuroMathComp) finalized their article [18] on the well-posedness of stochastic neural field equations within a rigorous framework.
- J. Inglis and E. Tanré together with F. Delarue and S. Rubenthaler (Univ. Nice – Sophia Antipolis) finalized their article [16] on the global solvability of a networked system of integrate-and-fire neurons proposed in the neuroscience literature.
- J. Inglis and E. Tanré together with F. Delarue and S. Rubenthaler (Univ. Nice – Sophia Antipolis) completed their study of the mean-field convergence of a highly discontinuous particle system modeling the behavior of a spiking network of neurons, based on the integrate-and-fire model [17]. Due to the highly singular nature of the system, it was convenient to work with a relatively unknown Skorohod topology.
- J. Inglis and D. Talay introduced in [38] a new model for a network of spiking neurons that attempted to address several criticisms of previously considered models. In particular the new model takes into account the role of the dendrites, and moreover includes non-homogeneous synaptic weights to describe the fact that not all neurons have the same effect on the others in the network. They were able to obtain mean-field convergence results, using new probabilistic arguments.
- A. Lejay have worked with G. Pichot (EPI Sage) on benchmarks for testing Monte Carlo methods to simulate particles in one-dimensional media, and applied this statistical methodology to four methods, including the exact method developed previously [45]. This work led also to empirical observations that should guide the design of new methods [24].
- S. Maire is working with the Bulgarian Academy of sciences on Monte Carlo algorithms for linear equations based on killed random walks. In a first work, with I. Dimov and J-M. Sellier [37], a new Monte Carlo method to solve linear systems of equations has been introduced. This method can either compute one component of the solution or all components simultaneously. In a second work, with Ivan Dimov and Rayna Georgieva, a new Monte Carlo method to solve Fredhom integral equations of the second kind is developed [36].
- D. Villemonais worked with P. Del Moral (Univ. Sydney) on the conditional ergodicity of time inhomogeneous diffusion processes [35]. They proved that, conditionally on non extinction, an elliptic time-inhomogeneous diffusion process forgets its initial distribution exponentially fast. An interacting particle scheme to numerically approximate the conditional distribution is also provided.
- D. Villemonais proved a Foster-Lyapunov type criterion which ensures the exponential ergodicity of a Fleming-Viot type particle system whose particles evolve as birth and death processes. The criterion also ensures the tightness of the sequence of empirical stationary distributions considered as a family of random measures. A numerical study of the speed of convergence of the particle system is also obtained under various settings [41].

6.1.2. Other works in progress

- M. Bossy and J-F. Jabir (University of Valparaíso) proved the validity of a particle approximation of a (simplified) Lagrangian Stochastic Model submitted to specular reflections at the boundary and satisfying the mean no-permeability condition. This work achieves to extend our previous study [43] to the multidimensional case.
- N. Champagnat and D. Villemonais obtained criterions for existence, uniqueness and exponential convergence in total variation of quasi-stationary distributions and Q-processes for general absorbed and killed diffusion processes. The criterion obtained is equivalent to the property that a diffusion on natural scale coming down from infinity has uniformly (w.r.t. the initial condition) bounded expectation at a fixed time t. A study of nearly critical cases allow to conjecture that this property is true for all diffusion processes on natural scale coming down from infinity. This work is currently being written.
- N. Champagnat and B. Henry worked on the long-time behavior of the frequency spectrum for the Splitting Tree models under the infinitly-many alleles model. They obtained, using a new method for computing the expectation of an integral with respect to a random measure, the asymptotic behavior

of the moments of the frequency spectrum. As an application, they derived the law of large number and a new central limit theorem for the frequency spectrum. This work is currently being written.

- J. Claisse defended his PhD. under the supervision of N. Champagnat and D. Talay on stochastic control of population dynamics. He completed a finite-horizon optimal control problem on branching–diffusion processes. He also created and studied a hybrid model of tumor growth emphasizing the role of acidity. Key therapeutic targets appear in the model to allow investigation of optimal treatment problems.
- J. Claisse and D. Talay in collaboration with X. Tan (Univ. of Paris Dauphine) extended their previous work on a pseudo-Markov property enjoyed by the solutions of controlled stochastic differential equations and its application to the proof of the dynamic programming principle. A paper is being finished.
- M. Deaconu and O. Lupascu are working with L. Beznea (Bucharest, Romania) on a stochastic model for avalanche phenomena involving rupture properties that occur in the physical and deterministic models for snow avalanches. This approach is based on their recent results on fragmentation processes by stochastic differential equation and branching processes.
- M. Deaconu and O. Lupascu are working on a numerical probabilistic algorithm for an avalanche-type process. The originality of this approach is to use a coagulation/fragmentation model to describe the avalanche phenomenon. More precisely, they consider a particular fragmentation kernel which introduces "rupture-type" properties of deterministic models for snow avalanches.
- An important issue in neuroscience is the modelling of spike trains of a single neuron. In this context, the membrane potential of a neuron can be described by using a simple stochastic differential equation with periodic input, that is reset to a rest potential each time it hits a certain threshold. J. Inglis, A. Richard, D. Talay, and E. Tanré study how the law of these hitting times is affected when one changes the white noise (in the SDE) into a correlated noise. Practically, they use a fractional Brownian motion, and since the computation of the hitting times of such a non-Markovian, non-semimartingale process is still an open question, they rather try to compute the deviations from the white noise model. This is expected to give insights on the relevance of models with memory and long-range dependence.
- J. Inglis started a collaboration with B. Hambly and S. Ledger at the University of Oxford, in which interacting mean-field particle systems with common noise are being studied. Such systems are representative of systems of spiking neurons or portfolio defaults. In previous studies each particle was driven by a noise that was assumed independent from particle to particle (i.e. intrinsic noise). By considering a common driving noise in addition to the intrinsic noise, it is possible to model the fact that the environment in which the particles live is also noisy. This leads to the study of a new type of conditional McKean-Vlasov equation.
- J. Inglis, in collaboration with J. Maclaurin (EPI NEUROMATHCOMP) and W. Stannat (Berlin), has begun working on a new framework to understand the effect of noise on neural field equations. Deterministic neural field equations exhibit traveling wave solutions, and so the effect of noise on these solutions is of great interest. The idea is to decompose the solution into various components, which allow one to see directly how the noise affects the solution in the direction of the moving wave front. In particular, the goal is to reconcile mathematically the previous works of P. Bressloff and W. Stannat on the same subject, and to obtain a large deviation principle.
- J. Inglis and D. Talay are in the process of studying the emergence of spatio-temporal noise starting from microscopic models of neuron conductance.
- A. Lejay continued his collaboration with S. Torres (Universidad de Valapraíso, Chile) and E. Mordecki (Universidad de la República, Uruguay) on the estimation of the parameter of the Skew Brownian motion. This work is related to the modelling of diffusion processes in media with interfaces and has potential applications in many domains, such as population ecology.

- Together with R. Rebolledo (Pontificia Universidad Católica, Santiago, Chile), A. Lejay continued his review work on the mathematical modelling of the Wave Energy Converter Called the Oscillating water column, within the framework of the CIRIC project.
- A. Lejay continued his work on the Snapping out Brownian motion to perform numerical tests for the computation of the mean residence time in a diffusive medium with semi-permeable membranes, such as the one encountered in the mathematical modelling of diffusion Magnetic Resonance Imaging.
- A. Lejay continued his collaboration with L. Coutin (Universté Paul Sabatier, Toulouse) on the sensitivity of rough linear differential equations, by providing general results on the derivatives of the solution of rough differential equations with respect to parameters or the starting point.
- S. Niklitschek Soto and D. Talay completed their stochastic analysis of diffraction parabolic PDEs with general discontinuous coefficients in the multidimensional case.
- P. Guiraud (University of Valparaíso) and E. Tanré study the effect of noise in the phenomenon of spontaneous synchronisation in a network of connected leaky integrate-and-fire neurons. They detail cases in which the phenomenon of synchronization persists in a noisy environment, cases in which noise permits to accelerate synchronization, and cases in which noise permits to observe synchronization while the noiseless model does not show synchronization. (Math Amsud program SIN)
- O. Faugeras (EPI NEUROMATHCOMP) and E. Tanré worked on an extension of [44] to a context of several populations of homogeneous neurons. They study the limit mean field equation of the membrane potential as the number of neurons increases in a network with correlated synaptic weights. A paper is in preparation.
- C. Graham (CMAP, Ecole polytechnique) and D. Talay are writing the second volume of their series published by Springer on the Mathematical Foundations of Stochastic Simulations.
- In collaboration with N. Touzi (CMAP, Ecole polytechnique), D. Talay is studying stochastic differential equations involving local times with stochastic weights, and extensions of classical notions of viscosity solutions to PDEs whose differential operator has discontinuous coefficients and transmission boundary conditions.

6.2. Financial Mathematics

Participants: Mireille Bossy, Nicolas Champagnat, Madalina Deaconu, Antoine Lejay, Khaled Salhi, Denis Talay, Etienne Tanré.

6.2.1. Published works and preprints

- In collaboration with N. Maïzi (CMA - Mines Paristech) and O. Pourtallier (COPRIN team, Inria Sophia Antipolis - Méditerranée), M. Bossy studied the existence of a Nash equilibrium between electricity producers selling their production on an electricity market and buying CO2 emission allowances on an auction carbon market. The producers' strategies integrate the coupling of the two markets via the cost functions of the electricity production. The authors set out the set of Nash equilibria on the electricity market, that constitutes an equivalence class (same prices and market shares) from which they exhibit a dominant strategy. On the coupled markets, given a specific carbon market design (in terms of penalty level and allowances), they compute the bounds of the interval where carbon prices (derived from the previous dominant strategy) evolve. They specify the properties of the associated equilibria (see [30] and [14]).
- In their article [40], N. Champagnat, M. Deaconu, A. Lejay and K. Salhi have constructed a regime switching model for estimating the Value-at-Risk. This model classifies the states in crisis and steady regimes and constructs a mixture of power laws as a model for returns of financial assets.

- In collaboration with V. Reutenauer and C. Michel (CA-CIB), D. Talay and E. Tanré worked on a model in financial mathematics including bid-ask spread cost. They study the optimal strategy to hedge an interest rate swap that pays a fixed rate against a floating rate. They present a methodology using a stochastic gradient algorithm to optimize strategies. A paper has been submitted [39].

6.2.2. Other works in progress

- In collaboration with J. Bion-Nadal (Ecole Polytechnique and CNRS), D. Talay pursued the study of a new calibration methodology based on dynamical risk measures and stochastic control PDEs.

TYREX Project-Team

6. New Results

6.1. Automated Refactoring for Size Reduction of CSS Style Sheets

Cascading Style Sheets (CSS) is a standard language for stylizing and formatting web documents [17]. Its role in web user experience becomes increasingly important. However, CSS files tend to be designed from a result-driven point of view, without much attention devoted to the CSS file structure as long as it produces the desired results. Furthermore, the rendering intended in the browser is often checked and debugged with a document instance. Style sheets normally apply to a set of documents, therefore modifications added while focusing on a particular instance might affect other documents of the set.

We present a first prototype and a new CSS semantic analyzer and optimizer that is capable of automatically detecting and removing redundant property declarations and rules. We build on earlier work on tree logics to locate redundancies due to the semantics of selectors and properties. Existing purely syntactic CSS optimizers can be used in conjunction with our tool, for performing complementary (and orthogonal) size reduction, toward the common goal of providing smaller and cleaner CSS files. We have been able to detect large numbers of unnecessary property declarations in complex web pages; and we have also found mistakes in the style sheets of some of the most popular web sites. The number of safe modifications can easily grow as more components of CSS are supported and more features are implemented, such as property inheritance, translation of pseudo-classes into query languages, analysis of media queries, merging of equivalent selectors or containment involving grouped selectors.

6.2. Equipping IDEs with XML-Path Reasoning Capabilities

One of the challenges in Web development is to achieve a good level of quality in terms of code size and runtime performance for popular domain-specific languages such as XQuery, XSLT, and XML Schema. We developed an IDE augmented with static detection of inconsistent XPath expressions that assists the programmer with simplifying development and debugging of any application involving XPath expressions [12]. The tool is based on newly developed formal verification techniques based on expressive modal logics, which are now efficient enough to be used in the process of software development. We applied this to a full XQuery compiler for which we introduced an analysis for identifying and eliminating dead code automatically.

6.3. XQuery and Static Typing: Tackling the Problem of Backward Axes

XQuery is a functional language dedicated to XML data querying and manipulation. As opposed to other W3C-standardized languages for XML (e.g. XSLT), it has been intended to feature strong static typing. Currently, however, some expressions of the language cannot be statically typed with any precision. This is due to a discrepancy between the semantics of the language and its type algebra: namely, the values of the language are (possibly inner) tree nodes, which may have siblings and ancestors in the data. The types on the other hand are regular tree types, as usual in the XML world: they describe sets of trees. The type associated to a node then corresponds to the subtree whose root is that node and contains no information about the rest of the data. This makes navigation expressions using “backward axes,” which return e.g. the siblings of a node, impossible to type.

We show how to handle this discrepancy by improving the type system. We describe a logic-based language of extended types able to represent inner tree nodes and show how it can dramatically increase the precision of typing for navigation expressions. We describe how inclusion between these extended types and the classical regular tree types can be decided, allowing a hybrid system combining both type languages. The result is a net increase in precision of typing [20].

6.4. A Core Calculus for XQuery 3.0: Combining Navigational and Pattern Matching Approaches

XML processing languages can be classified according to whether they extract XML data by paths or patterns. The strengths of one category correspond to the weaknesses of the other. In this work, we propose to bridge the gap between these two classes by considering two languages, one in each class: XQuery (for path-based extraction) and CDuce (for pattern-based extraction). To this end, we extend CDuce so as it can be seen as a succinct core λ-calculus that captures XQuery 3.0. The extensions we consider essentially allow CDuce to implement XPath-like navigational expressions by pattern matching and precisely type them. The elaboration of XQuery 3.0 into the extended CDuce provides a formal semantics and a sound static type system for XQuery 3.0 programs [18].

6.5. Session Types as Generic Process Types

Behavioural type systems ensure more than the usual safety guarantees of static analysis [15]. They are based on the idea of “types-as-processes”, providing dedicated type algebras for particular properties, ranging from protocol compatibility to race-freedom, lock-freedom, or even responsiveness. Two successful, although rather different, approaches, are session types and process types. The former allows to specify and verify (distributed) communication protocols using specific type (proof) systems; the latter allows to infer from a system specification a process abstraction on which it is simpler to verify properties, using a generic type (proof) system. What is the relationship between these approaches? Can the generic one subsume the specific one? At what price? And can the former be used as a compiler for the latter? This work is a step towards answers to such questions. Concretely, we have defined a stepwise encoding of a pi-calculus with sessions and session types (the system of Gay and Hole) into a pi-calculus with process types (the Generic Type System of Igarashi and Kobayashi). We encode session type environments, polarities (which distinguish session channels end-points), and labelled sums. We show forward and reverse operational correspondences for the encodings, as well as typing correspondences. To faithfully encode session subtyping in process types subtyping, one needs to add to the target language record constructors and new subtyping rules. This work shows how the programming convenience of session types as protocol abstractions can be combined with the simplicity and power of the pi-calculus, taking advantage in particular of the framework provided by the Generic Type System.

6.6. Personal Shopping and Navigator System for Visually Impaired People

We have developed a a personal assistant and navigator system for visually impaired people [14]. This system has been built using a set of domain specific languages based on XML such as OpenStreetMap extended for Augmented Reality. It demonstrate how partially sighted people could be aided by the technology in performing an ordinary activity, like going to a mall and moving inside it to find a specific product. We propose an Android application that integrates Pedestrian Dead Reckoning and Computer Vision algorithms, using an off-the-shelf Smartphone connected to a Smart-watch. The detection, recognition and pose estimation of specific objects or features in the scene derive an estimate of user location with sub-meter accuracy when combined with a hardware-sensor pedometer. The proposed prototype interfaces with a user by means of Augmented Reality, exploring a variety of sensorial modalities other than just visual overlay, namely audio and haptic modalities, to create a seamless immersive user experience. The interface and interaction of the preliminary platform have been studied through specific evaluation methods. The feedback gathered will be taken into consideration to further improve the proposed system.

URBANET Team

6. New Results

6.1. Highlights of the Year

Two scientific results can be distinguished in UrbaNet activity this year. First of all, the work did in collaboration with Orange Labs during the PhD thesis of O. Erdene-Ochir (defended in 2013) led to a patent [38] related to routing in wireless sensor networks under resiliency constraints.

A second important result is represented by the book chapter "Wireless Access Networks for Smart Cities" [31], a common contribution of all the permanent members of the team. We hope that this chapter will become the reference on wireless networking within the new and dynamic smart cities community.

6.2. Characterizing and measuring urban networks

Participants: R. Domga Komguem, M. Fiore, D. Naboulsi, P. Raveneau, R. Stanica, F. Valois

6.2.1. Collection and Analysis of Mobile Phone Data

Cellular communications are undergoing significant evolutions in order to accommodate the load generated by increasingly pervasive smart mobile devices. At the same time, recent generations of mobile phones, embedding a wide variety of sensors, have fostered the development of open sensing applications, such as network quality or weather forecast applications.

In this sense, we contributed with a novel privacy-preserving mobile data collection platform [21], leveraging the dynamic deployment of crowdsourcing tasks across a pouplation of mobile phones.

Using such data, or other datasets coming from network operators, we can propose dynamic access network mechanisms that adapt to customers' demands. To that end, one must be able to process large amount of mobile traffic data and outline the network utilization in an automated manner. In [28], we propose a framework to analyze broad sets of Call Detail Records (CDRs) so as to define categories of mobile call profiles and classify network usages accordingly. We evaluated our framework on a CDR dataset including more than 300 million calls recorded in an urban area over 5 months. We showed how our approach allows to classify similar network usage profiles and to tell apart normal and outlying call behaviors.

6.2.2. Generation and Analysis of Vehicular Mobility Datasets

The surge in vehicular network research has led, over the last few years, to the proposal of countless network solutions specifically designed for vehicular environments. A vast majority of such solutions has been evaluated by means of simulation, since experimental and analytical approaches are often impractical and intractable, respectively. The reliability of the simulative evaluation is thus paramount to the performance analysis of vehicular networks, and the first distinctive feature that has to be properly accounted for is the mobility of vehicles, i.e., network nodes. Notwithstanding the improvements that vehicular mobility modeling has undergone over the last decade, no vehicular mobility dataset was publicly available that captures both the macroscopic and microscopic dynamics of road traffic over a large urban region.

In [12], we present a realistic synthetic dataset, covering 24 hours of car traffic in a 400-km2 region around the city of Ko¨ln, in Germany. We describe the generation process and outline how the dataset improves the traces currently employed for the simulative evaluation of vehicular networks. We also show the potential impact that such a comprehensive mobility dataset has on the network protocol performance analysis, demonstrating how incomplete representations of vehicular mobility may result in over-optimistic network connectivity and protocol performance.

Moreover, using a similar methodology we contribute to the ongoing effort to define such mobility scenarios by introducing a second set of traces for vehicular network simulation, this time focusing on a highway environment. Our traces are derived from high-resolution real-world traffic counts, and describe the road traffic on two highways around Madrid, Spain, at several hours of different working days. We provide a thorough discussion of the real-world data underlying our study, and of the synthetic trace generation process [20] [35] [29]. Finally, we assess the potential impact of our dataset on networking studies, by characterizing the connectivity of vehicular networks built on the different traces. Our results underscore the dramatic impact that relatively small communication range variations have on the network. Also, they unveil previously unknown temporal dynamics of the topology of highway vehicular networks, and identify their causes.

6.2.3. Characterizing Novel Wireless Networks for Urban Intelligent Transportation Solutions

Vehicular networks are not the only contribution communication technologies can bring in the field of Intelligent Transportation Systems. Two other examples have been studied this year in the team.

The first example is related to traffic light control in an urban environment [17]. A traffic light controller takes as input an estimation of the number of vehicles entering the intersection and produces as output a light plan, with the objective to reduce the traffic jam. The quality of the input traffic estimation is a key consideration on the performance of the traffic light controller. The advent of Wireless Sensor Networks, with their relatively low deployment and operation price, led to the development of several sensor-based architectures for intersection monitoring. We show in this work that the solutions proposed in the literature are unrealistic in terms of communication possibilities and that they do not allow a measure of the vehicular queue length at a lane level. Based on extensive experimental results, we propose an energy efficient, low cost and lightweight multi-hop wireless sensor network architecture to measure with a good accuracy the vehicle queue length, in order to have a more precise vision of traffic at the intersection.

On a second example, these last years have witnessed the rise of the smart cities and several mechanisms to render the cities more sustainable and more energy-efficient. Among all different aspects, a noteworthy one is urban bike development. Besides the growing enthusiast provoked by bicycles and the benefit for health they bring, there still exists some reluctance in using bikes because of safety, road state, weather, etc. To counter-balance these feelings, there is a need to better understand bicycle users habits, path, road utilization rate in order to improve the bicycle path quality. In this perspective, in [25], we propose to deploy a set of mobile sensors on bicycles to gather this different data and to exploit them to make the bike easier and make people want to ride bicycles more often. Such a network will also be useful for several entities like city authorities for road maintenance and deployment, doctors and environment authorities, etc. Based on such a framework, we propose a first basis model that help to dimension the network infrastructure and the kind of data to be real time gathered from bikes. More specifically, we present a theoretical model that computes the quantity of data a bike will be able to send along a travel and the quantity of data a base station should be able to absorb. We have based our study on real data to provide first numerical results and be able to draw some preliminary conclusions and open new research directions.

6.3. Technology specific solutions

Participants: I. Augé-Blum, W. Bechkit, J. Cui, A. Mouradian, T. Lin, H. Rivano, R. Stanica, F. Valois

6.3.1. Medium Access Control in Wireless Sensor Networks

Protocols developed during the last years for Wireless Sensor Networks (WSNs) are mainly focused on energy efficiency and autonomous mechanisms (e.g. self-organization, self-configuration, etc.). Nevertheless, with new WSN applications, new QoS requirements appear, such as time constraints. Real-time applications require the packets to be delivered before a known time bound which depends on the application requirements. We particularly focus on applications which consist in alarms sent to the sink node. We propose Real-Time X-layer Protocol (RTXP) [8], a real-time communication protocol. RTXP is a MAC and routing real-time communication protocol that is not centralized, but instead relies only on local information. To the best of our knowledge, it is the first real-time protocol for WSNs using an opportunistic routing scheme in order to increase the packet delivery ratio. In the paper above, we describe the protocol mechanisms. We give theoretical bounds on the end-to-end delay and the capacity of the protocol. Intensive simulation results confirm the theoretical predictions and allow to compare RTXP with a real-time scheduled solution. RTXP is also simulated under harsh radio channel, in which case the radio link introduces probabilistic behavior. Nevertheless, we show that RTXP performs better than a non-deterministic solution. It thus advocates for the usefulness of designing real-time (deterministic) protocols even for highly unreliable networks such as WSNs.

Continuing on the idea of WSN applications with strict temporal constraints, these critical applications require correct behavior, reliability, and, of course, the respect of time constraints. Otherwise, if they fail, consequences on human life and the environment could be catastrophic. For this reason, we argue that the WSN protocols used in these applications must be formally verified. Unfortunately the radio link is unreliable, it is thus difficult to give hard guarantees on the temporal behavior of the protocols (on wired systems the link error probability is very low, so they are considered reliable). Indeed, in WSN a message may experience a very high number of retransmissions. The temporal guarantee has thus to be given with a probability that it is achieved. This probability must meet the requirements of the application. Network protocols have been successfully verified on a given network topology without taking into account unreliable links. Nevertheless, the probabilistic nature of radio links may change the topology (links which appear and disappear). Thus, instead of a single topology we have a set of possible topologies, each topology having a probability to exist. In this paper, we propose a method that produces the set of topologies, checks the property on every topology, and gives the probability that the property is verified. This technique is independent from the verification technique, i.e. each topology can be verified using any formal method which can give a "yes" or "no" answer to the question: "Does the model of the protocol respect the property?". In [27], we apply this method on the previously proposed f-MAC protocol, a real-time medium access protocol for WSNs. We use UPPAAL model checker as verification tool, and we perform simulations to observe the difference between average and worst case behaviors.

One WSN application gaining a lot of importance in the team in the last few years targets Intelligent Transportation Systems (ITS), as also explained in the previous section. In this ITS field, parking sensor networks are rapidly deploying around the world and are also regarded as one of the first implemented urban services in smart cities. To provide the best network performance in this context, the MAC protocol shall be adaptive enough in order to satisfy the traffic intensity and variation of parking sensors. In this sense, in [24] and [36], we compare the performance of two off-the-shelf medium access control protocols on two different kinds of traffic models, and then evaluate their application-end information delay and energy consumption while varying traffic parameters and network density. From the simulation results, we highlight some limits induced by network density and occurrence frequency of event-driven applications. When it comes to real-time urban services, a protocol selection shall be taken into account - even dynamically - with a special attention to the energy-delay trade-off. In a follow-up study [23], we use real world data, more precisely the heavy-tailed parking and vacant time models from the SmartSantander platform, and then we apply the traffic model in the simulation with four different kinds of MAC protocols, that is, contention-based, schedule-based and two hybrid versions of these. The result shows that the packet inter-arrival time is no longer heavy-tailed while collecting a group of parking sensors, and then choosing an appropriate MAC protocol highly depends on the network configuration. Also, the information delay is bounded by traffic and MAC parameters which are important criteria while the timely message is required.

6.3.2. Routing in Wireless Sensor Networks

Routing represents another major challenging issue in WSN, because of the application diversity and energy efficiency constraints. Gradient broadcast routing is a robust scheme for data gathering in WSNs. At each hop, the sender broadcasts the packet to its neighbors and one or more nodes among its neighbors closer to the sink forward it. As long as a node has at least one neighbor with a smaller hop-count, it can route packets. Nevertheless, nodes can disappear because of energy depletion, hardware failure, etc. In this case, it cannot be ensured that a packet reaches the sink. Usually this issue is addressed by updating the gradient with a periodical flooding. Nevertheless, it consumes an important amount of energy, moreover, parts of the network may not need to be updated. In [26], we propose GRABUP (GRAdient Broadcast UPdate), a traffic-based gradient maintenance algorithm which updates the gradient thanks to the data packets. We simulate the proposition and compare it with the classic gradient broadcast routing.

Another specific application that we target is smart metering, which heavily rely on the communication network for efficient data gathering, thus eliminating manual meter reading. Smart electronic devices are deployed in open, unattended and possibly hostile environment such as consumer's home and office areas, making them particularly vulnerable to physical attacks. Resilience is needed in this case to mitigate such inherent vulnerabilities and risks related to security and reliability. In [18], a general overview of the resilience including definition, metric and resilient techniques relevant for smart metering is presented. A quantitative metric, visual and meaningful, based on the graphical representation is adopted to compare routing protocols in the sense of resilience against active insider attacks. Five well-known routing protocols from the main categories have been studied through simulations and their resilience is evaluated according to the given metric. Resilient techniques introduced to these protocols have enhanced significantly the resilience against attacks providing route diversification.

6.3.3. Other Research Issues Related to Wireless Sensor Networks

Important features of WSNs, such as low battery consumption, changing topology awareness, open environment, non reliable radio links, raise other research issues than classical MAC and routing problems. For example, in [32], we investigate the benefits of Network Coding in WSN, especially with respect to resiliency. We have seen in our previous work that resiliency could be described as a multi dimensional metric, taking parameters such as Average Delivery Ratio, Delay Efficiency, Energy Efficiency, Average Throughput and Delivery Fairness into account. Resiliency can then be graphically represented as a kiviat diagram created by the previous weighted parameters. In order to introduce these metrics, previous works have been leaded on the Random Gradient Based Routing, which proved good resiliency in malicious environment. We look for seeing the improvements in term of resiliency, when adding network coding in the Random Gradient Based Routing with malicious nodes.

Another challenge is represented by the deployment of sensor nodes, which can take into account the impact of multiple parameters. For example, temperature variations have a significant effect on low power WSNs as wireless communication links drastically deteriorate when temperature increases. A reliable deployment should take temperature into account to avoid network connectivity problems resulting from poor wireless links when temperature increases. A good deployment needs also to adapt its operation and save resources when temperature decreases and wireless links improve. Taking into account the probabilistic nature of the wireless communication channel, we develop [4] a mathematical model that provides the most energy efficient deployment in function of temperature without compromising the correct operation of the network by preserving both connectivity and coverage. We use our model to design three temperature-aware algorithms that seek to save energy (i) by putting some nodes in hibernate mode as in the SO (Stop-Operate) algorithm, or (ii) by using transmission power control as in PC (Power-Control), or (iii) by doing both techniques as in SOPC (Stop-Operate Power-Control). All proposed algorithms are fully distributed and solely rely on temperature readings without any information exchange between neighbors, which makes them low overhead and robust. Our results identify the optimal operation of each algorithm and show that a significant amount of energy can be saved by taking temperature into account.

Finally, the notion of Shared Risk Link Groups (SRLG) captures survivability issues when a set of links of a network may fail simultaneously, such as a WSN where link conditions are extremely dynamic. The theory of survivable network design relies on basic combinatorial objects that are rather easy to compute in the classical graph models: shortest paths, minimum cuts, or pairs of disjoint paths. In the SRLG context, the optimization criterion for these objects is no longer the number of edges they use, but the number of SRLGs involved. Unfortunately, computing these combinatorial objects is NP-hard and hard to approximate with this objective in general. Nevertheless some objects can be computed in polynomial time when the SRLGs satisfy certain structural properties of locality which correspond to practical ones, namely the star property (all links affected by a given SRLG are incident to a unique node, for example a battery depleted sensor) and the span property (the links affected by a given SRLG form a connected component of the network). The star property is defined in a multi-colored model where a link can be affected by several SRLGs while the span property is defined only in a mono-colored model where a link can be affected by at most one SRLG. In [33], we extend these notions to characterize new cases in which these optimization problems can be solved in polynomial time or are fixed parameter tractable. We also investigate on the computational impact of the transformation from the multi-colored model to the mono-colored one. Experimental results are presented to validate the proposed algorithms and principles.

6.3.4. Data Aggregation and Gathering

In the data gathering problem, a particular network node, the base station or the sink, aims at receiving messages from some other network nodes. In [5], we model this network as a graph, and we consider that, at each step, a node can send one message to one of its neighbors (such an action is called a call). However, a node cannot send and receive a message during the same step. Moreover, the communication is subject to interference constraints, more precisely, two calls interfere in a step, if one sender is at distance below a certain threshold from the other receiver. Given a graph with a base station and a set of nodes having some messages, the goal of the gathering problem is to compute a schedule of calls for the base station to receive all messages as fast as possible, i.e., minimizing the number of steps (called makespan). The gathering problem is equivalent in this case to the personalized broadcasting problem where the base station has to send messages to some nodes in the graph, with same transmission constraints. We focus on the gathering and personalized broadcasting problem in grids (regular networks, with nodes deployed in a grid-like shape, e.g. parking or intersection monitoring WSNs). Moreover, we consider the non-buffering model: when a node receives a message at some step, it must transmit it during the next step. In this setting, though the problem of determining the complexity of computing the optimal makespan in a grid is still open, we present linear (in the number of messages) algorithms that compute optimal schedules for data gathering.

Data aggregation is a particular solution for the data gathering problem, which reduces the amount of data sent to the base station. In [16], we show that data aggregation can effectively reduce the energy consumption and improve the network capacity. Moreover, we present the state-of-the-art aggregation functions, including compressing-based and forecasting-based method; compressing-based aggregation focuses on compressing the data packets accompanied with transmitting based on spatial correlation, while forecasting aggregation tends to use mathematical models to fit the time series and predict the new value due to highly temporal correlation. We detail these two methods and characterize them respectively. We propose comparison between A-ARMA and Compressing Sensing, which are noteworthy examples of forecasting aggregation and compressing aggregation respectively.

6.3.5. Safety Vehicular Ad Hoc Networks

Vehicular ad hoc networks can play an important role in enhancing transportation efficiency and improving road safety. Therefore, direct vehicle-to-vehicle communications are considered as one of the main building blocks of a future Intelligent Transportation System. The success and availability of IEEE 802.11 radios made this technology the most probable choice for the medium access control layer in vehicular networks. However, IEEE 802.11 was originally designed in a wireless local area network context and it is not optimized for a dynamic, ad hoc vehicular scenario. In [11], we investigate the compatibility of the IEEE 802.11 medium access control protocol with the requirements of safety vehicular applications. As the protocols in this family are well-known for their scalability problems, we are especially interested in high density scenarios, quite frequent on today's roads. Using an analytical framework, we study the performance of the back-off mechanism and the role of the contention window on the control channel of a vehicular network. Based on these findings, we propose a reverse back-off mechanism, specifically designed with road safety applications in mind. Extensive simulations are carried out to prove the efficiency of the proposed enhancement scheme and to better understand the characteristics of vehicular communications.

One of the major roles of vehicular communication is the dissemination of information on the road in order to increase the awareness of the drivers. The facilities layer is a recently standardized component in the vehicular communication architecture, with an important role to play in the process of information dissemination. In [22], we propose facilities layer-based mechanisms for information propagation and we show they outperform classical network layer solutions. We also demonstrate that previous studies that do not consider the cohabitation of different types of safety messages on the vehicular control channel highly under-estimate the dissemination delay, which can lead to unrealistic assumptions in the design of safety applications.

6.4. Capillary solutions

Participants: M. Fiore, G. Gaillard, D. Naboulsi, H. Rivano, R. Stanica, F. Valois

6.4.1. Connected Vehicles

Bandwidth availability in the cellular backhaul is challenged by ever-increasing demand by mobile users. Vehicular users, in particular, are likely to retrieve large quantities of data, choking the cellular infrastructure along major thoroughfares and in urban areas. It is envisioned that alternative roadside network connectivity can play an important role in offloading the cellular infrastructure. We investigate [7] the effectiveness of vehicular networks in this task, considering that roadside units can exploit mobility prediction to decide which data they should fetch from the Internet and to schedule transmissions to vehicles. Rather than adopting a specific prediction scheme, we propose a fog-of-war model that allows us to express and account for different degrees of prediction accuracy in a simple, yet effective, manner. We show that our fog-of-war model can closely reproduce the prediction accuracy of Markovian techniques. We then provide a probabilistic graph-based representation of the system that includes the prediction information and lets us optimize content prefetching and transmission scheduling. Analytical and simulation results show that our approach to content downloading through vehicular networks can achieve a 70% offload of the cellular network.

Vehicles also produce large quantities of Floating Car Data (FCD), which consist of information generated by moving vehicles and uploaded to Internet-based control centers for processing and analysis. As upcoming mobile services based on or built for networked vehicles largely rely on uplink transfers of small-sized but high-frequency messages, FCD traffic is expected to become increasingly common in the next few years. Presently, FCD are managed through a traditional cellular network paradigm: however, the scalability of such a model is unclear in the face of massive FCD upload, involving large fractions of the vehicles over short time intervals. In [13], we explore the use of vehicle-to-vehicle (V2V) communication to partially relieve the cellular infrastructure from FCD traffic. Specifically, we study the performance boundaries of such a FCD offloading approach in presence of best- and worst-case data aggregation possibilities at vehicles. We show the gain that can be obtained by offloading FCD via vehicular communication, and propose a simple distributed heuristic that has nearly optimal performance under any FCD aggregation model.

We also advocate the use of a data shuttle service model to offload bulk transfers of delay-tolerant data from the Internet onto standard vehicles equipped with data storage capabilities [14]. We first propose an embedding algorithm that computes an offloading overlay on top of the road infrastructure. The goal is to simplify the representation of the road infrastructure as raw maps are too complex to handle. In this overlay, each logical link maps multiple stretches of road from the underlying road infrastructure. We formulate then the data transfer assignment problem as a novel linear programming model that determines the most appropriate logical paths in the offloading overlay for a data transfer request. We evaluate our proposal using actual road traffic counts in France. Numerical results show that we can satisfy weekly aggregate requests in the petabyte range while achieving cumulative bandwidth above 10 Gbps with a market share of 20% and only one terabyte of storage per vehicle.

6.4.2. Energy Consumption in Communication Networks

Providing high data rates with minimum energy consumption is a crucial challenge for next generation wireless networks. There are few papers in the literature which combine these two issues. The work we propose in [10] focuses on multi-hop wireless mesh networks using a MAC layer based on S-TDMA (Spatial Time Division Multiple Access). We develop an optimization framework based on linear programming to study the relationship between throughput and energy consumption. Our contributions are twofold. First, we formulate and solve, using column generation, a new MILP to compute offline energy-throughput tradeoff curve. We use a physical interference model where the nodes can perform continuous power control and can use a discrete set of data rates. Second, we highlight network engineering insights. We show, via numerical results, that power control and multirate functionalities allow optimal throughput to be reached, with lower energy consumption, using a mix of single hop and multihop routes.

Another strategy with regard to energy consumption is switching off some network nodes that are not carrying any data or control traffic. In [37], we tackle the problem of on-grid energy saving in cellular networks based on switch-on/off techniques for base stations and the usage of renewable energy. We aim to evaluate how much power can be saved in the network and dimension the renew able energy system according to the consumptions in real-world networks.

6.4.3. Service Level Agreements

The era of the Internet of Things (IoT) brings complexity and deployment costs in smart cities, particularly in WSNs. Utilities such as gas or water providers are keen on delegating the management of the communications to specialized firms, namely WSN Operators, that will share the WSN resource among their various clients. For this reason, in [34] we provide a guideline to write Service Level Agreements (SLAs) for IoT operation, borrowing a well studied concept from the web services domain. We extend the SLA definition with specific items that integrate the WSN constraints, and we facilitate the construction of complex metrics that express the performance of the WSN.

Furthermore. WSN operators will need a robust and reliable technology in order to guarantee QoS constraints in a wireless environment, as in the industrial world. IEEE 802.15.4e Time Slotted Channel Hopping (TSCH) is one good candidate. Moreover, the IETF experience in IP networks management is an important input for monitoring and QoS control over WSNs. In [19], we give formal guidelines for the implementation of a SLA architecture for operated WSNs. We distinguish the various formal algorithms that are necessary to operate a WSN according to SLAs, and determines which functional entities are necessarily technology-dependent. Detailed examples of such entities are developed in an IPv6 over IEEE 802.15.4e TSCH context, such as advocated in the IETF 6TiSCH Working Group.

VEGAS Project-Team

5. New Results

5.1. Non-linear computational geometry

Participants: Guillaume Moroz, Sylvain Lazard, Marc Pouget, Mohamed Yacine Bouzidi, Laurent Dupont, Olive Chakraborty, Rémi Imbach.

5.1.1. Solving bivariate systems and topology of plane algebraic curves

In the context of our algorithm Isotop for computing the topology of plane algebraic curves [3], we work on the problem of solving a system of two bivariate polynomials. We focus on the problem of computing a Rational Univariate Representation (RUR) of the solutions, that is, roughly speaking, a univariate polynomial and two rational functions which map the roots of the polynomial to the two coordinates of the solutions of the system. The PhD thesis of Yacine Bouzidi [10] presented several results on this theme obtained during the past three years.

Separating linear forms. We addressed the problem of computing a linear separating form of a system of two bivariate polynomials with integer coefficients, that is a linear combination of the variables that takes different values when evaluated at the distinct solutions of the system. The computation of such linear forms is at the core of most algorithms that solve algebraic systems by computing rational parameterizations of the solutions and this is the bottleneck of these algorithms in terms of worst-case bit complexity. We presented for this problem a new algorithm of worst-case bit complexity $\widetilde{O}_B(d^7 + d^6\tau)$ where d and τ denote respectively the maximum degree and bitsize of the input (and where $\widetilde{O}$ refers to the complexity where polylogarithmic factors are omitted and O_B refers to the bit complexity). This algorithm simplifies and decreases by a factor d the worst-case bit complexity of a previous algorithm we presented in 2013 [24]. Our new algorithm also yields, for this problem, a probabilistic Las-Vegas algorithm of expected bit complexity $\widetilde{O}_B(d^5 + d^4\tau)$. These results were presented at the *International Symposium on Symbolic and Algebraic Computation* in 2014 [15].

Solving bivariate systems & RURs. Given such a separating linear form, we presented an algorithm for computing a RUR with worst-case bit complexity in $\widetilde{O}_B(d^7 + d^6\tau)$ and a bound on the bitsize of its coefficients in $\widetilde{O}(d^2 + d\tau)$. We showed in addition that isolating boxes of the solutions of the system can be computed from the RUR with $\widetilde{O}_B(d^8 + d^7\tau)$ bit operations. Finally, we showed how a RUR can be used to evaluate the sign of a bivariate polynomial (of degree at most d and bitsize at most τ) at one real solution of the system in $\widetilde{O}_B(d^8 + d^7\tau)$ bit operations and at all the $\Theta(d^2)$ real solutions in only $O(d)$ times that for one solution. These results were submitted in 2013, revised in 2014 and will appear in 2015 in the *Journal of Symbolic Computation* [12].

This work is done in collaboration with Fabrice Rouillier (project-team Ouragan at Inria Paris-Rocquencourt).

5.1.2. Topology of the projection of a space curve

Let $\mathcal{C}$ be a real plane algebraic curve defined by the resultant of two polynomials (resp. by the discriminant of a polynomial). Geometrically, such a curve is the projection of the intersection of the surfaces $P(x,y,z) = Q(x,y,z) = 0$ (resp. $P(x,y,z) = \frac{\partial P}{\partial z}(x,y,z) = 0$), and generically its singularities are nodes (resp. nodes and ordinary cusps). State-of-the-art numerical algorithms cannot handle, in practice, the computation of the curve topology in non-trivial instances. The main challenge is to find numerical criteria that guarantee the existence and the uniqueness of a singularity inside a given box B, while ensuring that B does not contain any closed loop of $\mathcal{C}$. We solve this problem by providing a square deflation system that can be used to certify numerically whether B contains a singularity p. Then we introduce a numeric adaptive separation criterion based on interval arithmetic to ensure that the topology of $\mathcal{C}$ in B is homeomorphic to the local topology at p. The theoretical parts of these results are summarized in [18] and are to be combined with experimental data before submission to a journal.

5.1.3. Reflection through quadric mirror surfaces

We addressed the problem of finding the reflection point on quadric mirror surfaces, especially ellipsoid, paraboloid or hyperboloid of two sheets, of a light ray emanating from a 3D point source P_1 and going through another 3D point P_2, the camera center of projection. This is a classical problem known as Alhazen's problem dating from around 1000 A.D. and based on the work of Ptolemy around 150 A.D. [22], [27]. We proposed a new algorithm for this problem, using a characterization the reflection point as the tangential intersection point between the mirror and an ellipsoid with foci P_1 and P_2. The computation of this tangential intersection point is based on our algorithm for the computation of the intersection of quadrics [5], [21]. The implementation is in progress. This work is done in collaboration with Nuno Gonçalves, University of Coimbra (Portugal).

5.1.4. Describing the workspace of a manipulator

We studied the geometry of the solutions of the 3-R$\underline{\text{P}}$S parallel manipulator. In particular, a parallel manipulator usually has several solutions to the Direct Kinematic Problem. These solutions correspond to different *assembly modes*. A challenge is to find non-singular trajectories connecting different assembly modes. In the literature, this is done by encircling locally a cusp point of the discriminant variety in the joint space. In this work, we used tools from computer algebra to compute a partition of the work space in uniqueness domains. This allowed us to find global singularity-free trajectories reaching up to three assembly modes [16], [17].

5.2. Classical and probabilistic computational geometry

Participants: Sylvain Lazard, Marc Pouget.

5.2.1. Worst-case silhouette size of random polytopes

We studied from a probabilistic point of view the size of the silhouette of a polyhedron. While the silhouette size of a polyhedron with n vertices may be linear for some view points, several experimental and theoretical studies show a sublinear behavior for a wide range of constraints. The latest result on the subject proves a bound in $\Theta(\sqrt{n})$ on the size of the silhouette from a random view point of polyhedra of size n approximating non-convex surfaces in a reasonable way [7]. In this result, the polyhedron is considered given and the sizes of its silhouettes are averaged over all view points. We addressed the problem of bounding the worst-case size of the silhouette where the average is taken over a set of polyhedra. Namely, we consider random polytopes defined as the convex hull of a Poisson point process on a sphere in $\mathbb{R}^3$ such that its average number of points is n. We show that the expectation over all such random polytopes of the maximum size of their silhouettes viewed from infinity is $\Theta(\sqrt{n})$. This work, done in collaboration with Marc Glisse (Inria Geometrica) and Julien Michel (Université de Poitiers), was submitted this year to the *Journal of Computational Geometry* [28].

5.2.2. Recognizing shrinkable complexes is NP-complete

We say that a simplicial complex is shrinkable if there exists a sequence of admissible edge contractions that reduces the complex to a single vertex. We prove [14] that it is NP-complete to decide whether a (three-dimensional) simplicial complex is shrinkable. Along the way, we describe examples of contractible complexes which are not shrinkable. This work was done in collaboration with Dominique Attali (CNRS, Grenoble), Olivier Devillers and Marc Glisse (Inria Geometrica).

5.2.3. On point-sets that support planar graphs

A set of points is said universal if it supports a crossing-free drawing of any planar graph. For a planar graph with n vertices, if edges can be drawn as polylines with at most one bend, we exhibited universal point-sets of size n if the bend-points can be placed arbitrarily [26]. Furthermore, if the bend points are also required to be chosen in the universal set, we proved the existence of universal sets of subquadratic size, $O(n^2/\log n)$ [25]. More recently, we considered the setting in which graphs are drawn with curved edges. We proved that, surprisingly, there also exists a universal set of n points in the plane for which every n-vertex planar graph admits a planar drawing in which the edges are drawn as a circular arc [11].

VERIDIS Project-Team

6. New Results

6.1. Highlights of the Year

The veriT solver (section 5.1) participated in the SMT competition 2014, part of the Vienna Summer Of Logic Olympic Games, and received the gold medal for the SMT category.

6.2. Automated and Interactive Theorem Proving

Participants: Pascal Fontaine, Marek Košta, Manuel Lamotte Schubert, Stephan Merz, Thomas Sturm, Hernán Pablo Vanzetto, Uwe Waldmann, Daniel Wand, Christoph Weidenbach.

6.2.1. Combination of Satisfiability Procedures

Joint work with Christophe Ringeissen from the CASSIS project-team at Inria Nancy Grand Est, and Paula Chocron, a student at the University of Buenos Aires.

A satisfiability problem is often expressed in a combination of theories, and a natural approach consists in solving the problem by combining the satisfiability procedures available for the component theories. This is the purpose of the combination method introduced by Nelson and Oppen. However, in its initial presentation, the Nelson-Oppen combination method requires the theories to be signature-disjoint and stably infinite (to ensure the existence of an infinite model). The design of a generic combination method for non-disjoint unions of theories is clearly a hard task but it is worth exploring simple non-disjoint combinations that appear frequently in verification. An example is the case of shared sets, where sets are represented by unary predicates. Another example is the case of bridging functions between data structures and a target theory (e.g., a fragment of arithmetic).

The notion of gentle theory has been introduced in the last few years as one solution to go beyond the restriction of stable infiniteness, in the case of disjoint theories. In [26], [43], we adapt the notion of gentle theory to the non-disjoint combination of theories sharing only unary predicates, constants, and equality. As in the disjoint case, combining two theories, one of them being gentle, requires some minor assumptions on the other one. We show that major classes of theories, i.e., Loewenheim and Bernays-Schoenfinkel-Ramsey, satisfy the appropriate notion of gentleness introduced for this particular non-disjoint combination framework.

We have also considered particular non-disjoint unions of theories connected via bridging functions [27]. We present a combination procedure which is proved correct for the theory of absolutely free data structures. We consider the problem of adapting the combination procedure to obtain a satisfiability procedure for the standard interpretations of the data structure. We present an enumeration procedure that allows us to revisit the case of lists with length.

6.2.2. Type Synthesis for Set-Theoretic Proof Obligations

TLA$^+$ is a language for the formal specification of systems and algorithms whose first-order kernel is a variant of untyped Zermelo-Fraenkel set theory. Typical proof obligations that arise during the verification of TLA$^+$ specifications mix reasoning about sets, functions, arithmetic, tuples, and records. One of the challenges in designing an efficient encoding of TLA$^+$ proof obligations for the input languages of first-order automatic theorem provers or SMT solvers is to synthesize appropriate sorts for the terms appearing in a proof obligation, matching the type system of the target prover. We base this synthesis on the detection of “typing hypotheses” present in the proof obligations and then propagate this information throughout the entire formula. An initial type system [53] similar to the multi-sorted discipline underlying SMT-lib was not expressive enough for representing constraints such as domain conditions for function applications. We therefore developed a more expressive type system that includes dependent types, predicate types, and subtyping. Type synthesis in this system is no longer decidable but generates constraints that are submitted to SMT solvers during type reconstruction. When the constraints are valid, the translation of the formula becomes simpler, and checking it becomes correspondingly more efficient. When type construction does not succeed, the translator locally falls back to a sound, but inefficient "untyped" encoding where interpreted sorts such as integers are injected into the SMT sort representing TLA$^+$ values. In practice, this approach is found to behave significantly better than the original type system, and it extends easily to ATP proof backends. The results have been published at NFM 2014 [29], full details appear in Vanzetto's PhD thesis [11].

6.2.3. Syntactic Abstractions in First-Order Modal Logics

Joint work with Damien Doligez, Jael Kriener, Leslie Lamport, and Tomer Libal within the TLA$^+$ project at the MSR-Inria Joint Centre.

TLA$^+$ proofs mix first-order and temporal logics, and few (semi-)automatic proof tools support such languages. Moreover, natural deduction and sequent calculi, which are standard underpinnings for reasoning in first-order logic, do not extend smoothly to modal or temporal logics, due to the presence of implicit parameters designating the current point of evaluation. We design a syntactic abstraction method for obtaining pure first-order, respectively propositional modal or temporal, formulas from proof obligations in first-order modal or temporal logic, and prove the soundness of this "coalescing" technique. The resulting formulas can be passed to existing automatic provers or decision procedures for first-order logic (possibly with theory support), respectively for propositional modal and temporal logic. The method is complete for proving safety properties of specifications. This work was presented at the workshop on Automated Reasoning in Quantified Non-Classical Logic organized as part of Vienna Summer of Logic [33], and it has been implemented within TLAPS (section 5.2).

6.2.4. Satisfiability of Propositional Modal Logics via SMT Solving

Joint work with Carlos Areces from the National University of Córdoba, Argentina, and Clément Herouard, a student at ENS Rennes.

Modal logics extend classical propositional logic, and they are robustly decidable. Most existing decision procedures for modal logics are based on tableau constructions. Within our ongoing cooperation with members of the National University of Córdoba supported by the MEALS and MISMT projects (sections 8.3 and 8.4), we are investigating the design of decision procedures based on adding custom instantiation rules to standard SAT and SMT solvers. Our constructions build upon the well-known standard translation of modal logics to the guarded fragment of first-order logic. The idea is to let the solver maintain an abstraction of the quantified formulas, together with corresponding models. The abstraction is refined by lazily instantiating quantifiers, until either it is found to be unsatisfiable or no new instantiations need to be considered. We prove the soundness, completeness, and termination of the procedure for basic modal logic and several extensions. In particular, a smooth extension to hybrid logic makes use of the decision procedures for equality built into SMT solvers, yielding surprisingly simple correctness proofs. A presentation of this work has been accepted for publication in 2015.

6.2.5. First-Order Extensions to Support Higher-Order Reasoning

In contrast to higher-order logic, first-order logic provides automation and completeness. In order to increase the success rate of first-order proof procedures on translations of higher-order proof obligations, we developed two extensions to first-order logic:

- a polymorphic type system and
- declarations for inductive data types.

While the former can be seen as "just some kind of complication" to standard first-order reasoning procedures, the latter is an extension beyond first-order logic. We have shown how to keep first-order completeness in the presence of inductive data types while making use of the declarations for inferences and reductions that cannot be justified at the first-order level. The result is a superposition calculus extended with induction that shows impressive performance on standard benchmark sets when compared to existing approaches.

6.2.6. Decidability of First-Order Recursive Clause Sets

Recursion is a necessary source for first-order undecidability of clause sets. If there are no cyclic, i.e., recursive definitions of predicates in such a clause set, (ordered) resolution terminates, showing decidability. In this work we present the first characterization of recursive clause sets enabling non-constant function symbols and depth increasing clauses but still preserving decidability. For this class called BDI (Bounded Depth Increase) we present a specialized superposition calculus. This work has been published in the Journal of Logic and Computation [18].

6.2.7. Finite Quantification in Hierarchic Theorem Proving

Joint work with Peter Baumgartner and Joshua Bax from NICTA, Canberra, Australia.

Many applications of automated deduction require reasoning in first-order logic modulo background theories, in particular some form of integer arithmetic. A major unsolved research challenge is to design theorem provers that are "reasonably complete" even in the presence of free function symbols ranging into a background theory sort. For the case when all variables occurring below such function symbols are quantified over a finite subset of their domains, we have developed and implemented a non-naive decision procedure for extended theories on top of a black-box decision procedure for the EA-fragment of the background theory. In its core, it employs a model-guided instantiation strategy for obtaining pure background formulas that are equi-satisfiable with the original formula. Unlike traditional finite model finders, it avoids exhaustive instantiation and, hence, is expected to scale better with the size of the domains [25].

6.2.8. Developing Learning Strategies for Virtual Substitution

Joint work with Konstantin Korovin from the University of Manchester, UK.

During the past twenty years there have been a number of successful applications of real quantifier elimination methods based on virtual substitution. On the other hand, recently there has been considerable progress in (linear and non-linear) real arithmetic SMT-solving triggered by the idea to adopt from Boolean SAT-solving conflict analysis and learning techniques. In this work we do the first steps towards combining these two lines of research.

We consider linear real arithmetic SMT-solving. Inspired by related work for the Fourier-Motzkin method, we develop learning strategies for linear virtual substitution. For the first time, we formalize a virtual substitution-based quantifier elimination method—with and without our learning strategies—as formal calculi in the style of abstract DPLL [55]. We prove soundness and completeness for these calculi. Some standard linear programming benchmarks computed with an experimental implementation of our calculi show that the novel learning techniques combined with linear virtual substitution give rise to considerable speedups. Our implementation is part of the Reduce package Redlog, which is open-source and freely available.

This work gave rise to a publication at the CASC 2014 international workshop [28].

6.2.9. Efficient Cell Construction in Cylindrical Algebraic Decomposition

Joint work with Christopher W. Brown from the United States Naval Academy.

In their 2012 paper, de Moura and Jovanović [51] give a novel procedure for non-linear real SMT solving. The procedure uses DPLL-style techniques to search for a satisfying assignment. In case of a conflict, Cylindrical Algebraic Decomposition (CAD) is used to guide the search away from the conflicting state: On the basis of one conflicting point, the procedure learns to avoid in the future an entire CAD cell containing that point. The crucial part of this "model-based" approach is a function realizing this cell learning. Unfortunately, it is the main computational bottleneck of the whole procedure.

In 2014, we improved our cell learning procedure developed in 2013 by further theoretical investigation, which led to optimizations of the cell construction algorithm. This work gave rise to a publication in the Journal of Symbolic Computation [14].

In this publication we present an algorithm for the cell construction problem. Given a point and a set of polynomials, the algorithm constructs a single cylindrical cell containing the point, such that the polynomials are sign-invariant in the constructed cell. To represent a single cylindrical cell, a novel data structure is introduced. The algorithm, which is based on McCallum's projection operator, works with this representation and proceeds incrementally: First a cell representing the whole real space is constructed, and then refinement with respect to a single input polynomial is done to ensure the sign-invariance of this polynomial in the refined cell. We prove that our algorithm is correct and efficient in the following sense: First, the set of polynomials computed by our algorithm is a subset of the set constructed by the "model-based" approach, and second, the cell constructed by our algorithm is bigger than the cell constructed by the "model-based" approach.

6.3. Formal Methods for Developing Algorithms and Systems

Participants: Manamiary Andriamiarina, Jingshu Chen, Marie Duflot-Kremer, Dominique Méry, Stephan Merz.

6.3.1. Incremental Development of Distributed Algorithms

Joint work with Mohammed Mosbah and Mohammed Tounsi from the LABRI laboratory in Bordeaux, France, and with Neeraj Kumar Singh from the Department of Computing and Software, McMaster University, Hamilton, Canada.

The development of distributed algorithms and, more generally, of distributed systems, is a complex, delicate, and challenging process. The approach based on refinement helps to gain formality by using a proof assistant, and proposes to apply a design methodology that starts from the most abstract model and leads, in an incremental way, to the most concrete model, for producing a distributed solution. Our work helps formalizing pre-existing algorithms, developing new algorithms, as well as developing models for distributed systems.

Our research was initially supported by the ANR project RIMEL (see http://rimel.loria.fr). More concretely, we aim at an integration of the correct-by-construction refinement-based approach into the *local computation* programming model underlying the VISIDIA toolkit developed at LABRI for designing distributed algorithms expressed as a set of rewriting rules over graph structures.

In particular, we show how state-based models can be developed for specific problems [22] and how they can be simply reused by controlling the composition of state-based models through the refinement relationship. Traditionally, distributed algorithms are supposed to run on a fixed network, whereas we consider a network with a changing topology.

The contribution is related to the development of proof-based patterns providing effective help to the developer of formal models of applications [24], [12], [42]. Our patterns simplify the development of distributed systems using refinement and temporal logic.

6.3.2. Modeling Medical Devices

Formal modelling techniques and tools [30] have attained sufficient maturity for formalizing highly critical systems in view of improving their quality and reliability, and the development of such methods has attracted the interest of industrial partners and academic research institutions. Building high quality and zero-defect medical software-based devices is a particular domain where formal modelling techniques can be applied effectively. Medical devices are very prone to showing unexpected system behaviour in operation when traditional methods are used for system testing. Device-related problems have been responsible for a large number of serious injuries. Officials of the US Food and Drug Administration (FDA) found that many deaths and injuries related to these devices are caused by flaws in product design and engineering. Cardiac pacemakers and implantable cardioverter-defibrillators (ICDs) are among the most critical medical devices and require closed-loop modelling (integrated system and environment modelling) for verification purposes before obtaining a certificate from the certification bodies.

Clinical guidelines systematically assist practitioners in providing appropriate health care in specific clinical circumstances. Today, a significant number of guidelines and protocols are lacking in quality. Indeed, ambiguity and incompleteness are likely anomalies in medical practice. The analysis of guidelines using formal methods is a promising approach for improving them.

In [32], we give the semantics of refinement diagrams that are used in a refinement-based methodology for complex medical systems design, which possesses all the required key features. A refinement-based approach relying on formal verification, model validation using a model-checker, and refinement charts is proposed in this methodology for designing a high-confidence medical device. We show the effectiveness of this methodology for the design of a cardiac pacemaker system. Moreover, we organized a Dagstuhl seminar on the Pacemaker Challenge [20].

6.3.3. Analysis of Real-Time Concurrent Programs

Joint work with Nadezhda Baklanova, Jan-Georg Smaus, Wilmer Ricciotti, and Martin Strecker at IRIT Toulouse, France, and master student Jorge Ibarra Delgado, funded by the Airbus Foundation (see also section 7.1).

We investigate techniques for the formal verification of multi-threaded real-time programs. We assume that programs contain annotations that indicate the times for executing basic blocks, and that these annotations are enforced by the execution platform. Inspired by Safety-Critical Java [49], our partners in Toulouse developed a formal semantics for a fragment of Java in Isabelle/HOL. We designed techniques for formally ensuring the absence of concurrent accesses to shared resources in bounded-length executions of such programs. Specifically, we generate constraints that characterize the possible execution orders of the program, and then invoke an SMT solver in order to verify that no execution violates precedence constraints that ensure absence of conflicts. In the case where such an execution exists, we obtain a trace that exhibits the access conflict. Our technique has been implemented prototypically, and appears to scale much better than a previous analysis based on an encoding of programs as timed automata. The results have been published at AVoCS 2014 [15].

During his internship within the first year of the Erasmus Mundus master program on Dependable Software Systems, Jorge Ibarra Delgado investigated the possibility of adapting the JOP toolset for Safety-Critical Java, and in particular its Worst-Case Execution Time (WCET) analyzer, for obtaining suitable annotations for basic blocks.

6.3.4. Bounding Message Length in Attacks Against Security Protocols

Joint work with Myrto Arapinis from the University of Glasgow, UK.

Security protocols are short programs that describe communication between two or more parties in order to achieve security goals. Despite the apparent simplicity of such protocols, their verification is a difficult problem and has been shown to be undecidable in general. This undecidability comes from the fact that the set of executions to be considered is of infinite depth (an infinite number of protocol sessions can be run) and infinitely branching (the intruder can generate an unbounded number of distinct messages). Several attempts have been made to tackle each of these sources of undecidability. We have shown that, under a syntactic and reasonable condition of "well-formedness" on the protocol, we can get rid of the infinitely branching part. A journal version of this result, extending the set of security properties to which it is applicable and that particular includes authentication properties, has been published in Information and Computation [13].

6.3.5. Evaluating and Verifying Probabilistic Systems

Joint work with colleagues at ENS Cachan and University Paris Est Créteil.

Since its introduction in the 1980s, model checking has become a prominent technique for the verification of complex systems. The aim was to decide whether or not a system fulfills its specification. With the rise of probabilistic systems, new techniques have been designed to verify this new type of systems, and appropriate logics have been proposed to describe more subtle properties to be verified. However, some characteristics of such systems fall outside the scope of model checking. In particular, it is often of interest not to tell wether a property is satisfied but how well the system performs with respect to a certain measure. We have designed a statistical tool for tackling both performance and verification issues. Following several conference talks, two journal papers have been submitted. The first one presents the approach in details with a few illustrative applications. The second one focuses on biological applications, and more precisely the use of statistical model checking to detect and measure several indicators of oscillating biological systems.

VIRTUAL PLANTS Project-Team

5. New Results

5.1. Highlights of the Year

- Publication of a joint work with RDP at ENS-Lyon in the journal 'Nature'. In December 2013, a joint work on phyllotaxy with the RDP lab from ENS-Lyon was published online in the journal Nature [2]. This paper obtained the 2014 prize "la Recherche" in the biology category http://www.leprixlarecherche.com. Based on the analysis of phyllotaxis perturbations in mutants, this study sheds a new light on our interpretation of phyllotaxis, revisiting the standard model and suggesting that several fields based on auxin and cytokinin with different properties are required to provide robustness to phyllotaxis.
- To study and model morphogenesis, the team has been working in the last 8 years on modeling mechanical forces and deformations in tissues in collaboration with the UMR RDP at ENS-Lyon. This work has given rise to the development of a 3D computational framework to model the mechanics of 3D plant tissues during growth at cellular resolution and has been finalized this year with a publication in PLoS Comp. Biology (to appear in 2015). This framework makes it possible to construct models of meristem development, showing how the regulation of regional identities can lead to realistic shape development by dynamically modulating the mechanical properties of cells. It has been used also to study the influence of a specific signalling cascade (the ABP1-Kat1 signalling pathway) and its putative mechanical consequences on primordium initiation [25]. The expertize gained by our groups on physical models of plant tissue development has been wrapped up in a review paper [12].

5.2. Analysis of structures resulting from meristem activity

5.2.1. Acquisition and design of plant geometry

Participants: Frédéric Boudon, Christophe Pradal, Christophe Godin, Christian Fournier, Ibrahim Chedaddi, Mathilde Balduzzi, Julien Diener.

Virtual 3D model of plants are required in many areas of plant modeling. They can be used for instance to simulate physical interaction of real plant structures with their environment (ligth, rain, wind, pests, ...), to set up initial conditions of growth models or to assess their output against real data. In the past decade, methods have been developed to digitize plant architectures in 3D [61], [52]. These methods are based on direct measurements of position and shape of every plant organ in space. Although they provide accurate results, they are particularly time consuming. More rapid and automated methods are now required in order to collect plant architecture data of various types and sizes in a systematic way. In this aim, we explore the use of pictures, laser scanner and video.

- *Reconstruction of plant architecture from 3D laser scanner data.* (Chakkrit Preuksakarn, Mathilde Balduzzi, Frédéric Boudon, Christophe Godin, Pascal Ferraro [Labri, Bordeaux], Yassin Refahi)

  We investigate the possibility to use 3D laser scanners to automate plant digitizing. We are developing algorithms to reconstruct branching systems without leaves or foliage from scanner data or from scan simulated on plant mock-up obtained using different digitizing method.

For branching systems, a number of automatic methods were proposed in the literature to reconstruct plausible branching structures from laser scanner data. The question of their assessment and accuracy is however critical for further exploitation in biological applications. For this, we developed an evaluation pipeline to assess reconstruction accuracy. A laser scan database on which experts built reference reconstructions is used as a basis of the evaluation. A graphical editor has been developed to help experts to reconstruct semi automatically reference structures. The evaluation pipeline is given two plant structures and compares their organization. Similar elements are identified based on geometric criteria using an optimization algorithm. The organization of these elements is then compared and their similarity is quantified. Two indices of geometrical and structural similarities are defined, and automatic reconstructions can thus be compared to reference structures to assess their accuracy. The method is successful at capturing the variation of similarities between two structures as different levels of noise are introduced. A first comparative evaluation of the different methods of the literature has been designed and conducted. This work has been published in the special issue of Annals of Botany on FSPMs [16]. A procedure to automatically determine phyllotactic angles from scans of small plants has been added recently to the reconstruction pipeline.

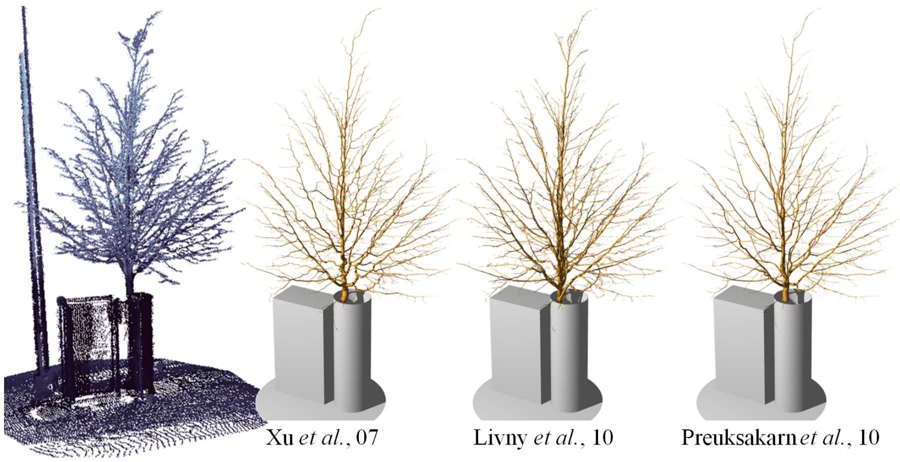


Figure 2. Comparison of different methods of the litterature to reconstruct plant architecture from laser scanner data. On the left, the original scan and on the rigth, the resulting reconstructions using three different methods. These reconstructions are quantitativelly assessed using our evaluation pipeline presented in [16]

In the context of the PhD of M. Balduzzi, we also investigated the reconstruction of tree foliage from 3D scans. Such elements are crucial to study the interaction of the plant with its environment. However, laser scans contain outliers in the silhouette of the scans that make the meshing of the pointset extremely difficult. New generation of laser scanners provide intensity of the laser reflected on the surface of scanned objects. This intensity depends on the distance to the object, its optical property and the incidence angle. A first work on this topic showed that after correcting the distance effect, the incidence angle can be deduced from the intensity. From this result, we developed a reconstruction pipeline using the scan intensities and based on Shape-From-Shading. Outliers being along the edge of the surface point cloud, we chose to develop a propagation SFS method initialized with points of the scans with high quality. We proved that surface with constant intensity are necessarily surfaces of constant slope or sand-pile surfaces. Using this result, a propagation method along iso-intensity regions was developed. These surfaces can then be sampled to provide a smooth point set without outliers. This work has been presented by M. Balduzzi for her thesis defense in november.

- *Reconstruction from video.* (Frédéric Boudon, Jerome Guenard [IRIT, Toulouse], Geraldine Morin [IRIT, Toulouse], Pierre Gurdjos [IRIT, Toulouse], Vincent Charvillat [IRIT, Toulouse])

  Even if mature computer vision techniques allow the reconstruction of challenging 3D objects from images, dedicated methods for generating 3D plant models must be devised due to the high complexity of plant topology. In collaboration with our colleagues from IRIT, Toulouse, we developed an analysis-by-synthesis method which generates 3D models of plants from both images and a priori knowledge of the plant species.

  Our method is based on a skeletonisation algorithm which allows to generate a possible skeleton from a foliage segmentation. Then, a 3D generative model, based on a parametric model of branching systems that takes into account botanical knowledge is built. This method extends previous works by constraining the resulting skeleton to follow hierarchical organization of natural branching structure. 3D models are then generated. 2D projections of the models can be compared with the original image to assess the visual accuracy of the reconstruction. We optimise the parameter values of the generative model based on the 2D projection criterion. Realistic results are obtained on different species of plants, in particular vineyards. This work has been presented at the Mathematical Methods for Curves and Surfaces conference and published in LNCS [42].

- *Reconstruction of virtual fruits from pictures.* (Ibrahim Chedaddi, Mik Cieslak, Nadia Bertin [Inra, Avignon], Frédéric Boudon, Christophe Godin, Michel Genard [Inra, Avignon], Christophe Goz-Bac [Université Montpellier 2])

  This research theme is supported by the Agropolis project MecaFruit3D.

  The aim of this work is to provide methods for generating fruit structure that can be integrated with models of fruit function. To this end, we used a modeling pipeline previously developped by Mik Cieslak in the OpenAlea platform. This methods involves two steps: (1) generating a 3D volumetric mesh representation of the entire fruit, and (2) generating a complex network of vasculature that is embedded within this mesh. Previous studies demonstrated the possibility to create species-specific models of fruit structure with relatively low effort [49]. We focus now on validating the vasculature networks by quantitatively comparing them to experimental data from the litterature.

  These physiological data will be combined with a mechanical model of fruit growth, to investigate the effects of fruit structure on quality (see section 5.4.2).

- *Reconstruction of root structures.* (Julien Diener, Fredéric Boudon, Christophe Pradal, Christophe Godin, Philippe Nacry [BPMP, INRA], Christophe Périn [AGAP, CIRAD], Anne Dievart [AGAP, CIRAD], Xavier Draye [UCL, Belgium])

  This research theme is supported by the Agropolis through the Rhizopolis project and by NUMEV.

  Similarly to aerial part of plants, new needs for automatic digitizing of root systems emerge. Most existing methods focus only on semi-automatic approaches. This does not support the high-throughput capabilities of acquisition systems. In the context of the RhizoScan project, we previously designed a prototype of an automatic image analysis pipeline to extract root system architecture of branching systems grown in Petri boxes. This pipeline provides i) a set of model based image segmentation method, ii) the extraction of a graph representation of the root system, and iii) a method to identify the root axes organization. This year, we improved and extended the pipeline in the following way:

  1. We integrated a validation step in the workflow based on the comparison method presented in [16].
  2. We participated in an international collaboration with the Université Catholique de Louvain (Belgium), the CPIB of the University of Notthingham (UK), the University of Vienna (Austria), the Jülich research center (Germany) and INRA, to develop a standard file format for root architecture. The resulting format (RSML) is described in a publication to appear in *Plant Physiology* in 2015 (details can be found at rootsystemml.github.io).

3. We have initiated a collaboration to integrate visual data mining methods developed by the Zenith team in order to improve the automation of the image analysis pipeline.
4. In general, the robustness of the pipeline has been improved. In particular, an optimization method has been designed to select the root axes hierarchy that respect specific botanical constraints.

5.2.2. Modeling the plant ontogenic programme

Participants: Christophe Godin, Yann Guédon, Jean-Baptiste Durand, Pierre Fernique, Christophe Pradal, Jean Peyhardi.

This research theme is supported by two PhD programmes.

The remarkable organization of plants at macroscopic scales may be used to infer particular aspects of meristem functioning. The fact that plants are made up of the repetition of many similar components at different scales, and the presence of morphological gradients, e.g. [46], [54], [55], [51], provides macroscopic evidence for the existence of regularities and identities in processes that drive meristem activity at microscopic scales. Different concepts have been proposed to explain these specific organizations such as "morphogenetic programme" [59], "age state" [50] or "physiological age" [48]. All these concepts state that meristem fate changes according to position within the plant structure and during its development. Even though these changes in meristem fate are specific to each species and lead to the differentiation of axes, general rules can be highlighted [50], [48]. Here we develop computational methods to decipher these rules.

- *Relating branching structure to the shoot properties* (Jean Peyhardi, Yann Guédon, Evelyne Coste [AGAP, AFEF team], Catherine Trottier [I3M], Yves Caraglio [AMAP], Pierre-Eric Lauri [AGAP, AFEF team])

  Shoot branching structures often take the form of a succession of homogeneous branching zones and have been analyzed using segmentation models such as hidden semi-Markov chains. Axillary meristem fates are influenced by local properties of the parent shoot such as for instance its growth rate or local curvature. The objective of this work, which is part of the PhD subject of Jean Peyhardi, is to develop statistical models that generalize hidden semi-Markov chains with the capability to incorporate explanatory variables that vary along the parent shoot (e.g. maximum growth rate of the leaf, surface of the leaf, length of the internode, local curvature of the parent shoot). More precisely, the simple multinomial distributions that represent the axillary productions observed in the different branching zones are replaced by multinomial generalized linear models (GLMs). Since the two classical categories of multinomial GLMs that correspond either to nominal or ordinal categorical response variables were not appropriate, we chose to develop a new family of multinomial GLMs called partitioned conditional GLMs [38] that enable to tackle hierarchically-structured categorical response variables. Typically, we need to distinguish different timing of branching (e.g. immediate shoot, one-year-delayed shoot and latent bud), different categories of offspring shoots (e.g. among one-year-delayed shoots, vegetative short shoot, vegetative long shoot and flowering shoot) and to specialize the explanatory variables for certain categories of offspring shoots (e.g. the growth of the parent shoot influence the immediate offspring shoots but not the one-year-delayed offspring shoots). The resulting integrative models are called semi-Markov switching partitioned conditional GLMs and have been applied to apple and pear tree branching structures.

- *Genetic determinisms of the alternation of flowering in apple tree progenies.* (Jean-Baptiste Durand, Jean Peyhardi, Baptiste Guitton [AGAP, AFEF team], Yan Holtz [AGAP, AFEF team] Catherine Trottier, Evelyne Costes, Yann Guédon)

  A first study was published to characterize genetic determinisms of the alternation of flowering in apple tree progenies. Data were collected at two scales: at whole tree scale (with annual time step) and a local scale (annual shoot, which corresponds to portions of stem that were grown during the same year). Two replications of each genotype were available.

Indices were proposed for early detection of alternation during the juvenile phase. They were based on a trend model and a quantification of the deviation amplitudes and dependency, with respect to the trend. This allows to quantify alternation from the yearly numbers of inflorescences at tree scale.

- *Identifying and characterizing patterns in tree-structured data* (Pierre Fernique, Jean-Baptiste Durand, Yann Guédon).

  In the context of Pierre Fernique's PhD (Montpellier 2 University and CIRAD), two complementary approaches were developed for analyzing patterns in tree-structured data:

  – multitype branching processes relying on local dependency properties for analyzing motifs.

  – multiple change-point models relying on long-term dependencies for segmenting trees in homogeneous zones.

  In multitype branching processes, the plant development is viewed as a demographic process, a parent entity of a given type generating child entities of different types (e.g. vegetative and flowering entities). Formally, the botanical entity properties are summarized as a categorical state variable. The number of child entities in each state is modeled through discrete multivariate distributions. Model selection procedures are necessary to specify parsimonious generation distributions. We developed an approach based on probabilistic graphical models to identify and exploit properties of conditional independence between numbers of children in different states, so as to simplify the specification of their joint distribution. The graph building stage was based on exploring the space of possible chain graph models, which required defining a notion of neighbourhood of these graphs [33]. To relax the strong constraints regarding dependencies induced by parametric distributions, mixture of graphical models were also considered [41]. Multitype branching processes were applied to the analysis of the patchiness pattern (consisting of canopies made of clumps of either vegetative or flowering growth units) in mango trees.

- *Simulating fruit tree phenology* (A.S. Briand, Frédéric Boudon, Frédéric Normand [CIRAD, HortSys, Réunion Island], Anaëlle Dambreville, Jean-Baptiste Durand, Pierre Fernique, Yann Guédon, Christophe Pradal, Pierre-Eric Lauri [AFEF team, AGAP])

  Mango is a tropical tree characterized by strong asynchronisms within and between trees. To study more precisely the interplay between the plant structural components, we built an integrative model to simulate the plant development based on the L-system formalism and GLM to model the dependencies between events. This model has been started during the internship of A. Jestin last year and has been continued during the visit of F. Boudon at the Hortsys lab in the Réunion island. The model has been extended this year to model growth and phenology of shoots and inflorescences (internship of A.S. Briand). For this, the sizes of the different organs is modelled by statistical laws estimated from measurements that depends on their positions in the architecture. The growth speed of organs is modulated by the temperature. This model will serve as a basis for further ecophysiological study in silico.

- *Integrative developmental growth stages of shoots* (Anaëlle Dambreville, Yann Guédon, Pierre-Eric Lauri [AFEF team, AGAP], Frédéric Normand [CIRAD, HortSys, Réunion Island])

  Plant growth, i.e. the increase of organ dimensions over time, and development, i.e. the change in plant structure, are often studied as two separate processes. However, there is structural and functional evidence that these two processes are strongly related. Our aim was to investigate the co-ordination between growth and development using mango trees, which have well-defined developmental stages. Developmental stages, determined in an expert way, and organ sizes, determined from objective measurements, were collected during the vegetative growth and flowering phases of two cultivars of mango. For a given cultivar and growth unit type (either vegetative or flowering), a multi-stage model based on absolute growth rate sequences deduced from the measurements was first built, and then growth stages deduced from the model were compared with hand-annotated developmental stages. Strong matches were obtained between both stages, leading to a consistent definition of integrative developmental growth stages. The growth stages highlighted growth asynchronisms between two topologically connected organs, namely the vegetative axis and its leaves. Integrative developmental growth stages emphasize that developmental stages are closely related to organ growth rates and can be interpreted in terms of the possible physiological processes (hydraulics, biomechanics and carbohydrate partitioning) underlying these stages.

- *Characterizing the successive flowering phases of strawberry in relation to genetic determinants* (Yann Guédon, Béatrice Denoyes [INRA, UMR BFP, Villenave d'Ornon], Justine Perrotte)

  Our aim was to characterize the successive flowering phases of perpetual flowering strawberry genotypes, which is of particular importance for better predicting fruit production. We applied multiple change-point models for the synchronous segmentation of the individuals of a given genotype in successive flowering phases. We identified two groups of genotypes that differ by the intensity of the flowering at the end of the flowering period. Using a genetic approach, we identified a locus controlling the flowering intensity at the end of the flowering period that likely explain these two groups of genotypes.

- *Self-nested structure of plants.* (Christophe Godin, Romain Azais, Farah Ben Naoum, Jean-Baptiste Durand, Alain Jean-Marie)

  In a previous work [6], we designed a method to compress tree structures and to quantify their degree of self-nestedness. This method is based on the detection of isomorphic subtrees in a given tree and on the construction of a DAG (Directed Acyclic Graph, equivalent to the original tree, where a given subtree class is represented only once (compression is based on the suppression of structural redundancies in the original tree). In the compressed graph, every node representing a particular subtree in the original tree has exactly the same height as its corresponding node in the original tree.

  The method proposed in [6] thus compresses a tree in width, but not in height. In a new work, we designed an extension of this compression method in which a tree is compressed in both width and height. The method is based on the detection of so-called *quasi-isomorphic paths* in a tree and on the compression of these paths in height. A paper describing the corresponding algorithms has been submitted recently for publication.

  The degree of self-nestedness is defined in [6] as the edit-distance between the considered tree structure and its nearest embedded self-nested version. Indeed, finding the nearest self-nested tree of a structure without more assumptions is conjectured to be an NP-complete or NP-hard problem. We thus designed a heuristic method based on interacting simulated annealing algorithms to tackle this difficult question. This procedure is also a keystone in a new topological clustering algorithm for trees that we propose. In addition, we obtain new theoretical results on the combinatorics of self-nested structures. The redaction of an article is currently in progress.

5.2.3. Analyzing the influence of the environment on the plant ontogenic programme

Participants: Jean-Baptiste Durand, Christian Fournier, Christophe Godin, Yann Guédon, Christophe Pradal, Jean Peyhardi, Pierre Fernique, Guillaume Garin.

This research theme is supported by three PhD programs.

The ontogenetic programme of a plant is actually sensitive to environmental changes. If, in particular cases, we can make the assumption that the environment is a fixed control variable (see section 5.2.2), in general the structure produced by meristem results from a tight interaction between the plant and its environment, throughout its lifetime. Based on observations, we thus aim to trace back to the different components of the growth (ontogenetic development and its modulation by the environment). This is made using two types of approaches. On the one hand, we develop a statistical approach in which stochastic models are augmented with additional time-varying explanatory variables that represent the environment variations. The design of estimation procedures for these models make it possible to separate the plant ontogenetic programme from its modulation by the environment. On the other hand, we build reactive models that make it possible to simulate in a mechanistic way the interaction between the plant development and its environment.

- *Influence of environment conditions and horticultural practices on the branching and axillary flowering structures of fruit tree shoots.* (Yann Guédon, Evelyne Costes [AFEF Team, AGAP], Ted DeJong [UC Davis], Claudia Negron [UC Davis]).

  In the context of a collaboration with Claudia Negron and Ted DeJong, we studied the influence of water availability [24] and pruning practices on the branching and axillary flowering structures of different categories of almond shoots. Stochastic models (hidden semi-Markov chains) were built for the branching and axillary flowering structures of different categories of almond shoots corresponding to different genetic backgrounds, levels of irrigation and pruning practices.

- *Analyzing growth components in trees.* (Yann Guédon, Yves Caraglio [AMAP], Olivier Taugourdeau [AMAP])

  We identified robust indicators that summarize the respective importance of ontogeny and environmental constraints (mainly related to light environment) in forest tree development. In this context, tree growth data correspond to the retrospective measurement of annual shoot characteristics (e.g. length, number of branches) along the main stem. We applied segmentation models to identify tree growth phases. These segmentation models, which are hidden semi-Markov chains, were compared with simple hidden Markov chains that correspond to the environment-driven development assumption. This statistical modelling approach was applied to both evergreen (Corsican pine and silver fir) and deciduous (sessile oak and Persian walnut) tree species growing in contrasted conditions ranging from managed forest stands to unmanaged understoreys. Growth phase duration distributions estimated within these segmentation models characterize the respective importance of ontogeny and environmental constraints in tree development at the population scale and have very contrasted characteristics in terms of shape and relative dispersion between ontogeny-driven and environment-driven tree development. These characteristics may change over tree life, reflecting changes in tree competition. Growth phase duration distributions summarize the joint trajectory of tree ontogeny and environment without requiring tree growth follow-up data for their estimation.

- *Analyzing fruit tree phenology in various climatic conditions from the shoot to the orchard scale* Yann Guédon, Juliano Dutra Schmitz [Universidade Federal de Pelotas, Brazil], Pierre-Eric Lauri [AFEF team, AGAP], Jean-Michel Legave [AFEF team, AGAP], Gustavo Malagui [Universidade Tecnológica Federal do Paraná]

  A first study was conducted at the shoot scale on four apple cultivars characterized by various chilling requirements and grown in two contrasting winter temperature conditions [20]. Our hypothesis was that shoot architecture is strongly affected by winter temperatures determining both the position and budburst of vegetative laterals with a lower effect on their outgrowth. A two-step approach was designed to quantify at the shoot scale i) the branching pattern and ii) two phenological stages of vegetative laterals, budburst and outgrowth. The branching pattern analysis combined various methods: branching sequence alignement and clustering, hidden semi-Markov chain and multiple change point model for segmenting branching sequences into homogeneous zones. A categorical variable, the branching zone, was built to summarize the lateral position along the shoot. It was integrated into the phenological analysis, based on a zero-inflated Poisson model, as a factor together with the cultivar and the winter temperature. We showed in this way that temperature had a main effect on the distribution of vegetative laterals along the shoot. It also strongly affected budburst, which was also affected by the cultivar and the branching zone. The outgrowth of the lateral was not significantly affected by temperature but was significantly affected by the cultivar and the branching zone. Furthermore, the delayed senescence and subsequent leaf persistence during winter, characterizing the apple tree in the mild winter temperature condition, had only a weak effect on the distribution of vegetative laterals and on budburst and lateral outgrowth. The actual shoot architecture and budburst result thus from an ordered sequence of events with a pivotal role of winter temperatures on the dormancy completion of individual lateral buds.

  A second study was conducted at the orchard scale. The time-course variation of dates of flowering stages was established for seventeen chronological sequences corresponding to various apple tree cropping areas in Europe (Belgium, France, Germany, Italy, Switzerland), north Africa (Morocco) and southern Brazil. Our aim was to characterize the relationship between flowering advances in fruit trees and global warming and to compare the northern and the southern hemisphere situations. We applied piecewise constant and linear homoscedastic models to these phenological series. The sudden advance of flowering dates detected at the end of the 1980s in the European locations can be explained by changes in rates for completion of heat requirements, essential to the development of floral primordia within buds. No effect of the global warming could be detected in the Brazilian flowering series and we only found a direct effect of the chilling temperature on the flowering date the same year (the colder the Austral winter, the earlier the flowering date).

- *Investigating how architectural development interfer with epidemics and epidemic control* (Christian Fournier, Corinne Robert [EGC], Guillaume Garin [ITK], Bruno Andrieu [EGC], Christophe Pradal)

  Sustainable agriculture requires the identification of new, environmentally responsible strategies of crop protection. Modelling of pathosystems can allow a better understanding of the major interactions inside these dynamic systems and may lead to innovative protection strategies. In particular, functional–structural plant models (FSPMs) have been identified as a means to optimize the use of architecture-related traits. A current limitation lies in the inherent complexity of this type of modelling, and thus the purpose of this paper is to provide a framework to both extend and simplify the modelling of pathosystems using FSPMs.

  Different entities and interactions occurring in pathosystems were formalized in a conceptual model [21]. A framework based on these concepts was then implemented within the open-source OpenAlea modelling platform, using the platform's general strategy of modelling plant–environment interactions and extending it to handle plant interactions with pathogens. New developments include a generic data structure for representing lesions and dispersal units, and a series of generic protocols to communicate with objects representing the canopy and its microenvironment in the OpenAlea platform. Another development is the addition of a library of elementary models involved in pathosystem modelling. Several plant and physical models are already available in OpenAlea and can be combined in models of pathosystems using this framework approach.

  Two contrasting pathosystems are implemented using the framework and illustrate its generic utility. Simulations demonstrate the framework's ability to simulate multiscaled interactions within pathosystems, and also show that models are modular components within the framework and can be extended. This is illustrated by testing the impact of canopy architectural traits on fungal dispersal. This study provides a framework for modelling a large number of pathosystems using FSPMs. This structure can accommodate both previously developed models for individual aspects of pathosystems and new ones. Complex models are dissassembled into separate *knowledge sources* originating from different specialist areas of expertise and these can be shared and reassembled into multidisciplinary models. The framework thus provides a beneficial tool for a potential diverse and dynamic research community.

5.3. Meristem functioning and development

In axis 2 work focuses on the creation of a *virtual meristem*, at cell resolution, able to integrate the recent results in developmental biology and to simulate the feedback loops between physiology and growth. The approach is subdivided into several sub-areas of research.

5.3.1. Data acquisition and design of meristem models

- *Improvement of the MARS-ALT pipeline robustness* Meristem, laser microscopy, image reconstruction, cell segmentation, automatic lineaging

  Participants: Léo Guignard, Christophe Godin, Christophe Pradal, Grégoire Malandain [Morpheme, Inria], Gaël Michelin [Morpheme, IPL Morphogenetics, Inria], Guillaume Baty, Sophie Ribes [IBC, UM2], Jan Traas [RDP, ENS], Patrick Lemaire [CRBM, CNRS], Yassin Refahi [RDP, ENS].

This research theme is supported by a PhD FRM grant, Jan Traas's ERC, Inria ADT programme and the Morphogenetics Inria Project Lab.

The MARS-ALT (Multi-Angles Registration and Segmentation - Automatic Lineage Tracking) software pipeline automatically performs a segmentation at cell resolution from 3D or 2D voxel images where the membranes/walls are marked (by a die for example) and makes it possible to follow the lineage of these cells through time [5]. This year, a new version of this pipeline has been developed that uses informations redundancy across the movies and biological knowledge on the segmented organism to constrain and therefore improve the segmentation and the tracking. To test the new pipeline, we used different acquisition protocols and different organisms (floral and apical meristems and the early stages of development of a marine animal *Phallusia mammillata*). The segmentation is corrected a posteriori to deal with imaging artifacts due to uncertainties of acquisition. The image data set on which we develop the methods consists of :

- *Arabidopsis thaliana* shoot apical meristem and primordia with around 6000 cells. The organ is captured from single angle every 4 hours during 2 or 3 days with a confocal microscope (Collaboration Sainsbury lab, Cambridge)
- *Arabidopsis thaliana* flower meristems with around 2000 cells. The organ is also captured from single angle with a confocal microscope (Collaboration RDP Lyon and Sainsbury lab)
- *Phallusia mammillata* embryos with from 32 cells to around 1000 cells. The organism is captured from four different angles every minute during 10 hours with a SPIM (Single Plane Illumination Microscope) (Collaboration CRBM Montpellier / EMBL Heidelberg). This work is developed in the context of the PhD work of Léo Guignard.

To our knowledgeIt is the first time that such high-resolution 4D digital tissues have been generated taking into account the cell shapes, opening the way to quantitative analysis of morphogenesis and tissue deformation at cell resolution.

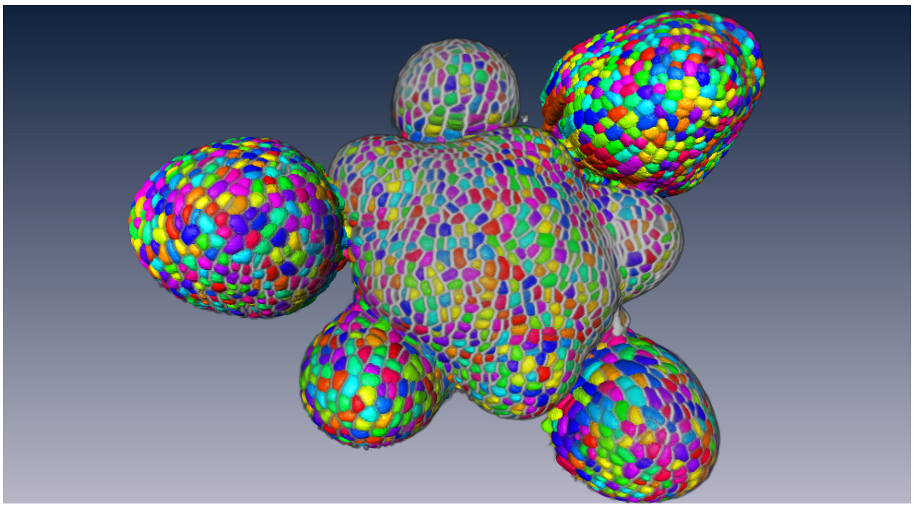


Figure 3. Superimposition of an automatic cell segmentation of an arabidopsis flower meristem using the new MARS pipeline with the original confocal image stack where the membranes are marked.

- *Creating mesh representation of cellular structures* (Guillaume Cerutti, Sohie Ribes, Christophe Godin)

Participants: Guillaume Cerutti, Sophie Ribes, Christophe Godin, Géraldine Brunoud [RDP, ENS], Carlos Galvan-Ampudia [RDP, ENS], Teva Vernoux [RDP, ENS], Yassin Refahi [RDP, ENS, Sainsbury Lab].

This research theme is supported the HFSP project Biosensors.

To produce a more efficient data structure accounting for the geometry of cellular tissues, we studied the problem of reconstructing a mesh representation of cells in a complex, multi-layered tissue structure, based either on membrane/wall images segmented using MARS or on nuclei images of shoot apical meristems. The construction of such mesh structures for plant tissues is currently a missing step in the existing image analysis pipelines. We developed a set of tools to build a triangular mesh surface representing the tissue in 3D, to evaluate the quality of the tissue reconstruction over objective aspects, to optimize a low-quality mesh simultaneously along several criteria, and to go towards a higher-scale representation pulling away from the cell resolution [31]. These methods are used in particular on nuclei images of shoot apical meristems of *Arabidopsis thaliana* to project hormonal information at cell-level on a continuous 3D tissue geometry. This work is carried out in the context of the post-doc of Guillaume Cerutti within the HFSP project BioSensors (Collaboration RDP Lyon).

These tools can produce light discrete representations of the cell interfaces that enables fast visualization, information projection, and quantitative analysis of the tissue, and give way to *in silico* physical and mechanical simulations on real-world data.

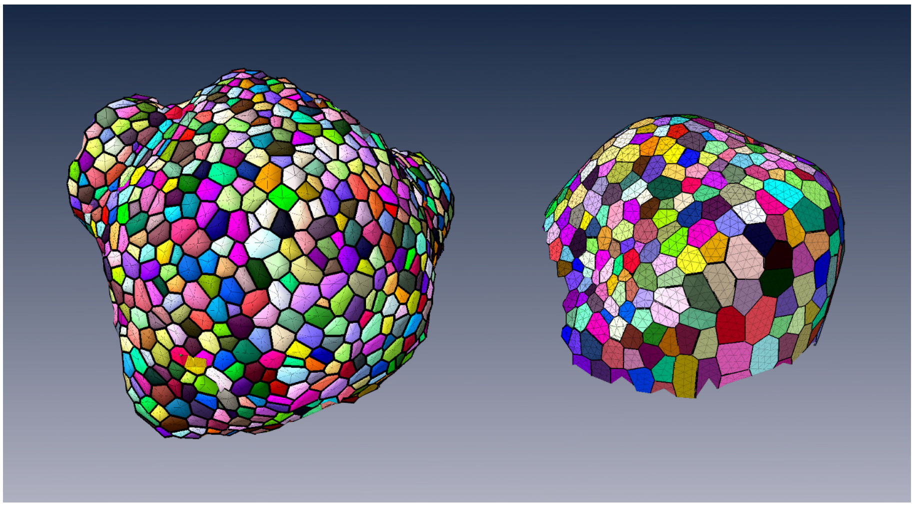


Figure 4. Triangular mesh representations of shoot apical meristem and flower meristem tissues obtained from MARS segmentations

- *Design of 3D digital atlases of tissue development*
  Participants: Sophie Ribes, Yassin Refahi [RDP, ENS, Sainsbury Lab], Guillaume Cerutti, Christophe Godin, Christophe Pradal, Christophe Pradal, Frédéric Boudon, Gregoire Malandain [RDP, ENS], Gaël Michelin [RDP, ENS], Guillaume Baty, Jan Traas [RDP, ENS], Teva Vernoux [RDP, ENS], Patrick Lemaire [CRBM, CNRS], Françoise Monéger [RDP, ENS].

  This research theme is supported the Inria Project Lab Morphogenetics, the ADT Mars-Alt and the HFSP project Biosensors.

  To organize the various genetic, physiological, physical, temporal and positional informations, we build a spatialized and dynamic database [56]. This database makes it possible to store all the collected information on a virtual 3D structure representing a typical organ. Each piece of information has to be located spatially and temporally in the database. Tools to visually retrieve and manipulate the information, quantitatively through space and time are being developed. For this, the 3D structure of a typical organ has been created at the different stages of development of the flower bud. This virtual structure contains spatial and temporal information on mean cell numbers, cell size, cell lineages, possible cell polarization (transporters, microtubules), and gene expression patterns. Such 3D digital atlas is mainly descriptive. However, like for classical databases, specific tools make it possible to explore the digital atlas according to main index keys, in particular spatial and temporal keys. Both a dedicated language and a 3D user interface are being designed to investigate and query the 3D virtual atlas. Current developments of this tool consist in using directly the segmented images produced from laser microscopy to build the atlas. To better represent the development of a biological population, a method to compute an "average" structure is investigated.

5.3.2. *Shape analysis of meristems*

Participants: Jonathan Legrand, Pierre Fernique, Frédéric Boudon, Yann Guédon, Christophe Godin, Pradeep Das [RDP, ENS], Arezki Boudaoud [RDP, ENS].

At cellular resolution, we studied the organization of cells in the meristems. The MARS-ALT pipeline provides rich spatio-temporal data sets for analyzing the development of meristems. A first step consisted of designing a dedicated graph for efficiently representing the spatial (adjacency between cells) and temporal (cell division) relationships between cells. Various variables can be attached either to the vertices (e.g. cell volume, inertia axes) or the edges (e.g. wall surface, distance between cell centroids). This graph may be augmented by new variables resulting from various spatial or temporal filtering (e.g. cell volumetric growth). Looking at homogeneous regions in the variable value space, cellular patterns can be identified. This work was developed in the context of the PhD of Jonathan Legrand with contributions of Pierre Fernique, another PhD student, that has been defended this year.

Considering the highly-structured nature of our data (time and space structuring) and the potential diversity and heterogeneity of possible cell descriptors, we developed two complementary approaches:

- A first one that favours the spatial structuring: In this approach, the cell neighbourhood and the cell descriptors are jointly taken into account in a clustering approach whose objective is to identify a small number of clusters corresponding to well-defined cell identities. Once the cells have been labelled using the clustering algorithm, cell generation distributions are estimated on the basis of the labelled lineage trees.
- A second one that favours the temporal structuring: In this approach, the data of interest are lineage forest and the only spatial structuring taken into account corresponds to siblings with respect to a given parent cell. In a first step, cell identities are inferred on the basis of the cell descriptors taking into account lineage relationships using hidden Markov tree models and the spatial regions that emerge from the cell identity labelling are then characterized. This second approach is supported by the fact that cell topology is only affected by division which makes highly relevant the local spatial information taken into account in this approach.

5.3.3. *Mechanical model*

Participants: Jean-Philippe Bernard, Olivier Ali, Christophe Godin, Benjamin Gilles, Frédéric Boudon, Ibrahim Cheddadi, Jan Traas [ENS-Lyon], Olivier Hamant [ENS-Lyon], Arezki Boudaoud [ENS-Lyon].

This research theme is supported by the Inria Project Lab Morphogenetics and the Jan Traas's ERC.

The rigid cell walls that surround plant cells are responsible for the acquisition of organ shapes. These walls are submitted to stresses due to cell turgor pressure. Wall deformation is caused by the turgor forces in the cell walls. Wall synthesis is triggered by these wall deformations when some specific threshold is exceeded. The final shape of the tissue integrates mechanically all the local deformations of each cell.

To quantify this growth process at the level of a multicellular tissue, we developed a model of growth that integrates mechanical forces development at cellular resolution. In this model, walls are characterized by their mechanical properties, e.g. elasticity, extensibility and anisotropy. For this, we used a tensorial approach to describe both tissue deformation and stresses. Deformations were decomposed into elementary transformations that can be related to underlying biological processes. However, we showed that the observed deformations does not map directly local growth instructions given by genes and physiology in each cell. Instead, the growth is a two-stage process where genes are specifying how cell walls should yield to mechanical stresses. In this way, different regions in the tissue with different cell identities can have different growth properties. The final shape of the tissue results from the integration of all these mechanical properties and stresses at organ level under the growth force due to turgor pressure at tissue scale.

A paper describing the mechanical model and its application to model primorium formation in the shoot apical meristem will appear in PLoS Computational Biology in 2015. We used this framework to investigate the influence of a specific signalling cascade (the ABP1- Kat1 signalling pathway) and its putative mechanical consequences on primordium initiation [25]. A review of the different mechanical concepts underlying plant morphogenesis has been carried out in [12].

In our first approach, the mechanical model rely on a finite element method (FEM) to describe the deformation of the tissue. In FEM, the tissue is represented by a mesh. The positions of the vertices at each time step are estimated from a linear system. If the tissue is big or if the mesh is fine, the linear system can be large and thus leads to computational overheads. An alternative way to classical FEM is to use a meshless method where the deformation of the tissue can be characterized by a linear combination of deformations of a finite and small set of frames. Because shape functions are no longer defined on each element but on the whole tissue, they have to be updated at each growth step by estimating a new rest configuration. With meshless method, the discretization of the system can be dynamically updated parsimoniously according to the precision required to model the emergence of shapes (PhD work of Jean-Philippe Bernard).

5.3.4. Gene regulatory networks: Design of a genetic model of inflorescence development.

Participants: Eugenio Azpeitia, Christophe Godin, François Parcy, Etienne Farcot.

This research theme is supported by the Inria Project Lab Morphogenetics.

Modeling gene activities within cells is of primary importance since cell identities correspond to stable combination of gene expression.

We studied the regulatory network that controls the flowering transition during morphogenesis. To overcome the network complexity and integrate this regulation during ontogenesis, we have developed a first model of the control of floral initiation by genes, and in particular the situation of cauliflower mutants, in which the repeatedly meristem fails in making a complete transition to the flower. Three different network models were done and validate. A first Boolean version, a second fuzzy logic and an ODEs models were studied. The models are able to correctly recover the gene steady states observed in the meristems during the flower transitions, the gene transitions and the mutant effects. Importantly, the model is able to explain the cauliflower mutants. This work couples models at different scales, since the gene regulatory network is used as a decision module in an L-system model of the inflorescence architecture. This mixed model has led us to make different hypotheses about gene interactions and hormonal regulation. First predictions about gene actors controling the passage to flower could be verified. Finally, some links between gene regulation and plant growth have been identified. These links can be experimentally tested which could lead to a first integrated picture of flower development could be reached in the context of Eugenio Azpeitia postdoc.

5.3.5. Model integration

Participants: Frédéric Boudon, Christophe Godin, Guillaume Baty, Jan Traas.

This research theme is supported by the Morphogenetics Inria Project Lab.

Our approach consists of building a programmable tissue which is able to accept different modeling components. This includes a central data structure representing the tissue in either 2-D or 3-D, which is able to grow in time, models of gene activity and regulation, models of signal exchange (physical and chemical) between cells and models of cell cycle (which includes cell division). For each modeling component, one or several approaches are investigated in depth, possibly at different temporal and spatial scales, using the data available from the partners (imaging, gene networks, and expression patterns). Approaches are compared and assessed on the same data. The objective of each submodel component will be to provide plugin components, corresponding to simplified versions of their models if necessary, that can be injected in the programmable tissue platform. This work is developed in collaboration with the RDP group at ENS-Lyon [58] and the CPIB group in Nottingham, UK [47].

One key aspect of our approach is the development of a computer platform dedicated to programming virtual tissue development. This platform will be used to carry out integration of the different models developed in this research axis. The platform is based on *OpenAlea*. In the past year, progress has been made in defining a generic tissue data structure that could be used in this platform. Currently, robust geometric operations such as division are implemented and tested. Moreover, a redesign of the structure based on more elaborated formalisms such as combinatorial maps is investigated.

5.4. Multi-scale models and analysis: from cells to plant architecture (and back)

5.4.1. Modeling water transport in roots

Participants: Mikaël Lucas [IRD], Christophe Pradal, Christophe Godin, Christophe Maurel [BPMP].

This research theme is supported by the ANR project HydroRoot.

A model of Arabidopsis thaliana root hydraulics at the cellular level was developped in the OpenAlea modeling platform. The model relies on the integration throughout root architecture of elementary hydraulic components. Each component integrates local radial and axial water flows. Axial hydraulic conductivity is calculated according to Poiseuille's law, based on local size of xylem vessels. Radial hydraulic conductivity is determined in part by aquaporin activity and was set constant throughout root architecture in the first model versions. In its current state, the model is parameterized using architectural, tissular and physiological data that were experimentally determined in the Aquaporin group at BPMP. The architectural reconstruction of the root system is based on a tridimensional multi-scale tree graph (MTG). The current model is capable of predicting the water flow that is transported by a root system in the standard experimental conditions used in the Aquaporin group. This model was used to perform sensitivity analyses and determine the respective contributions to root hydraulic dynamics of various biological parameters (axial and radial hydraulic conductivites, root architecture). One major finding is that the root hydraulic conductivity (Lpr) computed from the model is highly dependent on root architecture. This is due to the limiting role of axial (xylem) conductance, one feature that had been neglected in previous representations of root water transport. The radial hydraulic conductivity may primarily be limiting in conditions of Lpr inhibition, since its increase from values in control roots has marginal effects on Lpr. A new set of experimental data including root diameter repartitions in wild-type plants, and xylem vessel diameters in mutants with altered xylem morphology (irx3, esk1) will be used to implement the model. Root cell hydraulic conductivities will also be measured in these and aquaporin mutant phenotypes. Our aim is to check whether, based on anatomical and morphological data, the model can properly predict the radial hydraulic conductivity of these genotypes.

5.4.2. Mechanical modeling of fruit growth

Participants: Ibrahim Cheddadi [Inra, Avignon], Mik Cieslak [U. Calgary], Frédéric Boudon, Valentina Baldazzi [Inra, Avignon], Nadia Bertin [Inra, Avignon], Michel Genard [Inra, Avignon], Christophe Godin.

This research theme is supported by the Agropolis project MecaFruit3D.

Understanding the controlling factors of fruit quality development is challenging, because fruit quality results from the interplay between physical and physiological processes that are under the control of genes and the environment. Although process-based models have been used to make significant progress in understanding these factors, they ignored to a large extent the shape and internal structure of the fruit, as well as mechanical interactions between tissue parts that are essential to properly model growth.

To help characterizing the effects of fruit shape and internal structure on quality, the creation of a 3D virtual fruit model that integrates fruit structure and function with growth governed by environmental inputs is being investigated, combining two tools previously developed in the team: on the one hand, a modeling pipeline that creates a 3D volumetric mesh of the internal fruit structure, including vasculature (see section 3), and couples it with water and carbon transport; on the other hand, a mechanical description of the growth of plant tissues (see section 5.3.3): growth is related to the extension of the cell walls, which is triggered when the so-called turgor pressure inside the cells exceeds a given threshold. The global shape of the tissue integrates mechanically all the local deformations of each cell.

In order to couple these two aspects of plant growth, we describe how volume variations are constrained by fluxes of matter, and how these fluxes depend on mechanical and physiological parameters. The corresponding set of equations are resolved thanks to the SOFA finite elements software.

This approach will be applied to study tomato fruit. Once the model is calibrated and evaluated, our approach will be suitable for studying the effects of internal fruit heterogeneity and overall shape on fruit quality development.

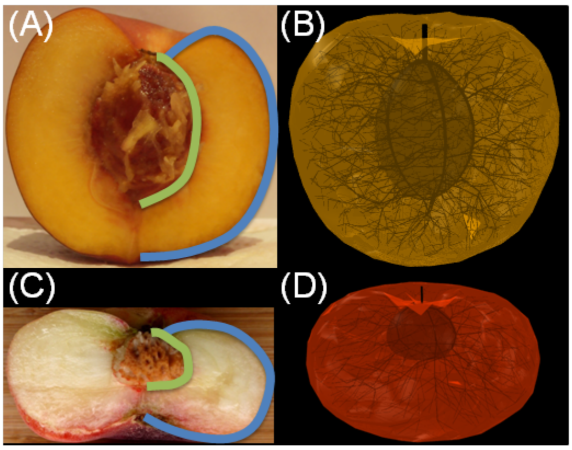


Figure 5. Virtual models of peaches reconstructed from images with simulated vasculatur to simulate carbon and water transport in the fruit

5.4.3. Analyzing root growth and branching

Participants: Beatriz Moreno Ortega, Sixtine Passot, Yann Guédon, Laurent Laplaze [IRD, DIADE], Mikaël Lucas [IRD, DIADE], Bertrand Muller [INRA, LEPSE].

This research theme is supported by two PhD programmes.

New 2D and 3D root phenotyping plateforms are emerging with associated image analysis toolbox (e.g. SmartRoot). The analysis of complex root phenotyping data is thus a new challenge in developmental biology.

We aim at developing a pipeline of methods for analyzing root systems at three scales:

1. tissular scale to identify and characterize the meristem, elongation and mature zones along a root using piecewise heteroscedastic linear models,
2. individual root scale to analyze the dynamics of root elongation,
3. root system scale to analyze the branching structure.

This pipeline of analysis methods will be applied to different species (maize, millet and *arabidopsis*) and for different biological objectives (study of genetic diversity for millet and of metabolic and hormonal controls of morphogenesis for maize).

5.4.4. Analyzing shoot and leaf elongation

Participants: Maryline Lièvre, Yann Guédon, Leo Guignard, Christine Granier [INRA, LEPSE].

This research theme is supported by one PhD programme.

This study is based on the observation of a lack of methods enabling the integrated analysis of the processes controlling the vegetative development in *Arabidopsis thaliana*. We developed a pipeline of analysis methods combining image analysis techniques and statistical models to integrate the measurements made at the leaf and shoot scales. Semi-Markov switching models were built for different genotypes, allowing a more thorough characterization of the studied mutants. These models validated the hypothesis that the rosette can be structured into successive developmental phases that could change depending on the genotype. They also highlighted the structuring role of the abaxial trichome trait, although the developmental phases cannot be explained entirely by this trait. We developed a second pipeline of analysis methods combining a semi-automatic method for segmenting leaf epidermis images based on the ilastik software, and the analysis of the obtained cell areas using a gamma mixture model whose parameters of gamma components are tied by a scaling rule. This model allowed us to estimate the distribution of the number of endocycles. We highlighted in this way that the mean number of endocycles changes drastically with leaf rank. Finally, we built a multi-scale model that integrates tissular, morphological, dynamical and dimension traits for each successive leaf along the shoot. This model gave us for the first time an integrative view of the development of the *Arabidopsis* rosette.

5.4.5. Analyzing perturbations in Arabidopsis thaliana phyllotaxis

Participants: Yassin Refahi, Fabrice Besnard, Yann Guédon, Christophe Godin, Etienne Farcot, Teva Vernoux [RDP, ENS].

This research theme has been supported by iSam, IBC and the Inria Project Lab Morphogenetics.

The geometric arrangement of lateral organs along plant stems, named phyllotaxis, shows a variety of striking patterns with remarkable regularities and symmetries. This has interested biologists, physicists, mathematicians and computer scientists for decades. These studies have lead to a commonly accepted standard interpretation of phyllotaxis that postulates that organs inhibit the formation of new organs in their vicinity. At a molecular scale, these inhibitory fields have been shown to result from the spatio-temporal distribution of the plant hormone auxin. This model theoretically explains a large part of the diversity of phyllotactic patterns observed in plants.

The cytokinin hormones are known to play a significant role in the regulation of phyllotaxis. Fabrice Besnard and Teva Vernoux realized that *Arabidopsis thalianaahp6* mutants, which are perturbed in the cytokinin signaling pathway, showed unusual chaotic perturbations of the phyllotaxis at macroscopic level.

In order to characterize these perturbations, we designed a pipeline of models and methods [60], [53] which relies of combinatorial and statistical techniques. Using this pipeline of methods, we have shown that the perturbation patterns in both wild-type and mutant plants can be explained by permutations in the order of insertion along the stem of 2 or 3 consecutive organs.The number of successive synchronized organs between two permutations reveals unexpected patterns that depend on the nature of the preceding permutation (2- or 3-permutation). We identified significant individual deviations of the level of baseline segments with reference to 137.5°, which confirms theoretical model predictions. Finally, we highlighted a marked relationship between permutation of organs and defects in the elongation of the internodes in between these organs.

We then looked at the origin of these permutations using confocal microscopy and realized that organs were in fact frequently co-initiated in the mutant, leading after development randomly in half of the cases to permutations. We concluded that the mutant is actually perturbed in the time between consecutive organ initiation (i.e. the plastochrone), while relative angular positions are not affected. After closer inspection, we realized that the mutated gene encodes a protein diffusing from the organs and creating a field around the organs that regulates the plastochrone. We could demonstrate that in the mutant, the absence of this field leads to co-initiations and subsequently to the observed permutations.

Altogether, this study sheds a new light on our interpretation of phyllotaxis, revisiting the standard model and suggesting that several fields based on auxin and cytokinin with different properties are required to provide robustness to phyllotaxis. An overview of this work has been published in the journal Nature [15].

5.4.6. A stochastic model of phyllotaxis

Participants: Yassin Refahi, Christophe Godin, Etienne Farcot, Teva Vernoux [RDP, ENS].

This research theme has been supported by IBC and the Inria Project Lab Morphogenetics.

To proceed further and find a mechanistic interpretation of these results, we are currently developing a stochastic extension of the standard model of phyllotaxis. We first analyzed the properties of the inhibitory fields created by the existing primordia on the initiation of new promordia, and concluded that the angular positions of organs are very robust to perturbations while plastochrons may be dramatically affected. This suggested that there exists a strong decoupling between space and time in the patterning process. To account for this observation, we modeled the perception of the initiation signal by cells using stochastic processes coupled with the intensity of inhibitory fields and showed that the observed permutation patterns emerge spontaneously from this purely local processes. A paper describing this model will be submitted for publication in 2015.

5.4.7. The role of auxin and sugar in rose bud outgrowth control

Participants: Jessica Bertheloot [INRA, Angers], Frédéric Boudon, Christophe Godin.

Auxin in the stem is known to be a key regulator of apical dominance. Over the last decades, many studies have been undertaken to understand its action mode, which is indirect because auxin in the main stem does not enter into the bud. Recently, apical dominance over basal buds in pea has been related to low sugar availability caused by high sugar demand of growing apical organs. Auxin and sugar are two signals regulating the entrance of bud into sustained growth in opposite ways. Recently it has been demonstrated that sugar effect on bud outgrowth was preceded by a modification of the hormonal network involved in auxin effect, which suggests that auxin and sugar pathways do interact in a non-trivial way. But auxin and sugar effects have been studied separately until now. In this work, we investigate what is the combined effect of sugar and auxin on bud outgrowth, and how they integrate to control bud entrance into sustained growth. For this, a series of experiments has been carried on a single-node cuttings of Rosa hybrida grown in vitro in which different combinations of sugar and auxin levels have been tested. A model the regulatory networks controling stem-bud molecular interaction is currently being investigated.

VISAGES Project-Team

6. New Results

6.1. Highlights of the Year

Dr Camille Maumet was awarded by the French Society of Magnetic Resonance in Biology and Medicine (SFRMBM) for her PhD Thesis on analysis of neuroimaging data including images from functional Magnetic Resonance Imaging (fMRI) and Arterial Spin Labeling http://www2.warwick.ac.uk/fac/sci/wmg/idh/idhnews/?tag=Neural+Engineering.

Dr Americ Stamm was awarded by the Univ. of Rennes I foundation as the best PhD thesis in Math, Computer Sciences and Electrical Engineering. This award is dedicated for the PhDs having the highest potential for innovation and technological transfer
;ses-de-la-fondationhttps://fondation.univ-rennes1.fr/les-prix-de-thèses-de-la-fondation.

6.2. Image Computing: Detection, Segmentation, Registration and Analysis

6.2.1. A Mathematical Framework for the Registration and Analysis of Multi-Fascicle Models for Population Studies of the Brain Microstructure

Participant: Olivier Commowick.

Diffusion tensor imaging (DTI) is unable to represent the diffusion signal arising from multiple crossing fascicles and freely diffusing water molecules. Generative models of the diffusion signal, such as multi-fascicle models, overcome this limitation by providing a parametric representation for the signal contribution of each population of water molecules. These models are of great interest in population studies to characterize and compare the brain microstructural properties. Central to population studies is the construction of an atlas and the registration of all subjects to it. However, the appropriate definition of registration and atlasing methods for multi-fascicle models have proven challenging. This paper proposes [24] a mathematical framework to register and analyze multi-fascicle models. Specifically, we define novel operators to achieve interpolation, smoothing and averaging of multi-fascicle models. We also define a novel similarity metric to spatially align multi-fascicle models. Our framework enables simultaneous comparisons of different microstructural properties that are confounded in conventional DTI. The framework is validated on multi-fascicle models from 24 healthy subjects and 38 patients with tuberous sclerosis complex, 10 of whom have autism. We demonstrate the use of the multi-fascicle models registration and analysis framework in a population study of autism spectrum disorder. This work was performed in close collaboration with the Children's Hospital Boston.

6.2.2. Longitudinal Intensity Normalization in Multiple Sclerosis Patients

Participants: Yogesh Karpate, Olivier Commowick, Christian Barillot, Gilles Edan.

In recent years, there have been many Multiple Sclerosis studies using longitudinal MR images to study and characterize the MS lesion patterns. The intensity of similar anatomical tissues in MR images is often different because of the variability of the acquisition process and different scanners. We proposed [29] a novel methodology for a longitudinal lesion analysis based on intensity standardization to minimize the inter-scan intensity difference. The intensity normalization maps parameters obtained using a robust Gaussian Mixture Model (GMM) estimation not affected by the presence of MS lesions. Experimental results demonstrated that our technique accurately performs the task of intensity standardization. We show consequently how the same technique can improve the results of longitudinal MS lesion detection.

6.2.3. Simultaneous Estimation of T1, T2 and B1 Maps From Relaxometry MR Sequences

Participants: Fang Cao, Olivier Commowick, Elise Bannier, Christian Barillot.

Interest in quantitative MRI and relaxometry imaging is rapidly increasing to enable the discovery of new MRI disease imaging biomarkers. While DESPOT1 is a robust method for rapid whole-brain voxelwise mapping of the longitudinal relaxation time (T1), the approach is inherently sensitive to inaccuracies in the transmitted flip angles, defined by the B1 inhomogeneity field, which become more severe at high field strengths (e.g., 3T). We have proposed [26] a new approach for simultaneously mapping the B1 field, M0 (proton density), T1 and T2 relaxation times based on regular fast T1 and T2 relaxometry sequences. The new method is based on the intrinsic correlation between the T1 and T2 relaxometry sequences to jointly estimate all maps. It requires no additional sequence for the B1 correction. We evaluated our proposed algorithm on simulated and in-vivo data at 3T, demonstrating its improved accuracy with respect to regular separate estimation methods.

6.2.4. Quantitative Relaxation Templates for the Human Brain at 3T

Participants: Fang Cao, Olivier Commowick, Christian Barillot.

Quantitative MRI (qMRI) templates of relaxation times and proton density can be of particular interest for dedicated clinical applications such as characterizing brain tissue abnormalities, as well as general research purposes. We have developed in [27] 3D qMRI statistical templates consisting of T1, T2, T2* and ρ* maps from the human brain at 3T. The qMRI templates were built from a population of 20 normal controls, for which individual maps were estimated in a robust manner, accounting for acquisition artifacts and expected relationships between the relaxometry parameters. For validation, we fed the qMRI templates into a realistic MRI simulator to synthesize MR-weighted images, and compared these images with the real MR acquisitions. High correlation coefficients (>0.80) show that the developed qMRI templates can be used as input dataset for MRI simulation community, which may be of great interest to clinical neuroscience field.

6.2.5. Myelin Water Fraction Imaging in Multiple Sclerosis patients

Participants: Olivier Commowick, Elise Bannier, Christian Barillot.

Multi-echo T2 relaxometry is a relevant imaging method for Myelin Water Fraction (MWF) quantification in the study of multiple sclerosis (MS). However, to ensure accurate estimation, a large number of echoes are still required that can drive to very long acquisitions. In practice, 32 echo times (TE) ranging from 10 ms to 320 ms and an echo spacing (ESP) of 10 ms are used. Analysis of the decay curve of the consecutive echoes allows the estimation of the T2 spectrum. The proposed approach makes use of recent spatial regularization methods for MWF estimation from clinically compatible acquisitions (typically 11 echoes acquired within 6 minutes with TE1=ESP=8.4 ms). The algorithms were evaluated on both synthetic and clinical data, illustrating the ability to compute accurate MWF maps from a low number of echoes. The 2 methods used a priori information as well as conventional and fast algorithm (NNLS), and a cross-validation strategy. Based on simulated and clinical data results, the nlsrNNLS estimation is more accurate and less penalizing than srNNLS. This regularization provides an efficient way to circumvent an ill-posed problem aspect, in particular with a reduced number of echoes for clinically acceptable acquisition times, allowing for accurate MWF estimation. This work, performed in the master internship of Lucas Soustelle, was accepted as a conference abstract at SFRMBM 2015, and is submitted to ISMRM.

6.3. Image processing on Diffusion Weighted Magnetic Resonance Imaging

6.3.1. Fast Identification of Optimal Fascicle Configurations from Standard Clinical Diffusion MRI Using Akaike Information Criterion

Participants: Olivier Commowick, Christian Barillot.

Analytic multi-compartment models have gained a tremendous popularity in the recent literature for studying the brain white matter microstructure from diffusion MRI. This class of models require the number of compartments to be known in advance. In the white matter however, several non-collinear bundles of axons, termed fascicles, often coexist in a same voxel. Determining the optimal fascicle configuration is a model selection problem. We have proposed [30], [33] a novel approach to identify such a configuration from clinical diffusion MRI where only few diffusion images can be acquired and time is of the essence. Starting from a set of fitted models with increasing number of fascicles, we use Akaike information criterion to estimate the probability of each candidate model to be the best Kullback-Leibler model. These probabilities are then used to average the different candidate models and output an MCM with optimal fascicle configuration. This strategy is fast and can be adapted to any multi- compartment model. We illustrate its implementation with the ball-and-stick model and show that we obtain better results on single-shell low angular resolution diffusion MRI, compared to the state-of-the-art automatic relevance detection method, in a shorter processing time.

6.3.2. Tracking the Cortico-Spinal Tract as a Multi-Modal Distribution of Streamlines from Local White Matter Microstructure Models

Participant: Olivier Commowick.

We have presented [31] a pipeline to reconstruct the corticospinal tract (CST) that connects the spinal cord to the motor cortex. The proposed method combines a new white matter microstructure model coined Diffusion Directions Imaging and a new tractography algorithm based on a particle filter adapted for approximating multi-modal distributions. In this paper, we put the computation time and accuracy of our pipeline to the test in the context of the MICCAI 2014 DTI challenge, which aims to provide fast and accurate reconstructions of the CST for presurgical planning of brain tumor extraction. These two key performance metrics are expected in such a situation where time is of the essence and the quality of the data is dependent on the patient's health condition and ability to cooperate. In no more than 1.5 hours per patient, we successfully provide accurate CSTs of 2 very collaborative patients who underwent a diffusion MRI protocol that included 69 diffusion-sensitizing gradients spread over 4 different shells ranging from b = 200 to b = 3000 s/mm^2.

6.3.3. Model selection improvement with non-central chi estimation of multi-compartment models

Participant: Olivier Commowick.

Diffusion images are known to be corrupted with a non-central chi (NCC)-distributed noise. There has been a number of proposed image denoising methods that account for this particular noise distribution. However, to the best of our knowledge, no study was performed to assess the influence of the noise model in the context of diffusion model estimation. In particular, multi-compartment models are an appealing class of models to describe the white matter microstructure but require the optimal number of compartments to be known a priori. Its estimation is no easy task since more complex models will always better fit the data, which is known as over-fitting. However, MCM estimation in the literature is performed assuming a Gaussian-distributed noise. We have shown in a preliminary study [32] that using the appropriate NCC distribution for modeling the noise model reduces significantly the over-fitting, which could be helpful for unraveling model selection issues and obtaining better model parameter estimates.

6.3.4. Symmetric Block-Matching Registration for the Distortion Correction of Echo-Planar Images

Participants: Renaud Hédouin, Olivier Commowick, Elise Bannier, Christian Barillot.

We have introduced a new approach to correct geometric and intensity distortion of Echo Planar Images (EPI) from images acquired with opposite phase encoding directions. A new symmetric block-matching registration algorithm has been developed for this purpose relying on new adapted transformations between blocks and a symmetric optimization scheme to ensure an opposite symmetric transformation. We present results of our algorithm showing its ability to robustly recover EPI distortion while obtaining sharper results than the popular TOPUP algorithm. This work was performed in close collaboration with the Children's hospital in Boston.

6.4. Medical Image Computing in Brain Pathologies

6.4.1. Adaptive Dictionary Learning for Competitive Classification of Multiple Sclerosis Lesions

Participants: Hrishikesh Deshpande, Pierre Maurel, Christian Barillot.

The manual delineation of Multiple Sclerosis (MS) lesions is a challenging task pertaining to the requirement of neurological experts and high intra- and inter-observer variability. It is also time consuming because large number of Magnetic Resonance (MR) image slices are needed to obtain 3-D information. Over the last years, various models combined with supervised or unsupervised classification methods have been proposed for segmentation of MS lesions using MR images. Recently, signal modeling using sparse representations (SR) has gained tremendous attention and is an area of active research. SR allows coding data as sparse linear combinations of the elements of over-complete dictionary and has led to interesting image recognition results. In this work, we have proposed to use a sparse representation and an adaptive dictionary learning paradigm to automatically classify Multiple Sclerosis (MS) lesions from MRI. In particular, we investigate the effects of learning dictionaries specific to the lesions and individual healthy brain tissues, which include White Matter (WM), Gray Matter (GM) and Cerebrospinal Fluid (CSF). The dictionary size plays a major role in data representation but it is an even more crucial element in the case of competitive classification. We present an approach that adapts the size of the dictionary for each class, depending on the complexity of the underlying data. The proposed algorithm is evaluated on 3-D multi-channel MR images demonstrating improved classification.

6.4.2. *Predictive Value of Imaging Markers at Multiple Sclerosis Disease Onset Based on Gadolinium- and USPIO- Enhanced MRI and Machine Learning*

Participants: Olivier Commowick, Jean-Christophe Ferré, Elise Bannier, Gilles Edan, Christian Barillot.

A novel characterization of Clinically Isolated Syndrome (CIS) patients according to lesion patterns has been proposed in [13]. More specifically, patients are classified according to the nature of inflammatory lesions patterns. It is expected that this characterization can infer new prospective figures from the earliest imaging signs of Multiple Sclerosis (MS), since it can provide a classification of different types of lesions across patients. The method is based on a two-tiered classification. Initially, the spatio-temporal lesion patterns are classified. The discovered lesion patterns are then used to characterize groups of patients. The patient groups are validated using statistical measures and by correlations at 24-month follow-up with hypointense lesion loads. The methodology identified 3 statistically significantly different clusters of lesion patterns showing p-values smaller than 0.01. Moreover, these patterns defined at baseline correlated with chronic hypointense lesion volumes by follow-up with an R^2 score of 0.90. The proposed methodology is capable of identifying three major different lesion patterns that are heterogeneously present in patients, allowing a patient classification using only two MRI scans. This finding may lead to more accurate prognosis and thus to more suitable treatments at early stage of MS.

6.4.3. *Robust detection of multiple sclerosis lesion from intensity-normalized multi-channel MRI*

Participants: Yogesh Karpate, Olivier Commowick, Christian Barillot.

Multiple sclerosis (MS) is a disease with heterogeneous evolution among the patients. Better understanding of the disease will lead to improved patient-adapted therapeutic strategies. We propose a novel paradigm to detect MS lesions based on a statistical framework which consists of detection based on differences between multi-channel MRI of patients and controls. This framework fused with intensity standardization was applied to the study of MS and highlighted the great interest of quantitative MRI measurements for a better characterization of MS. Experimental results demonstrate that our technique accurately detects significant differences in lesions consequently improving the results of MS lesion detection. This work has been accepted to SPIE Medical Imaging 2015.

6.4.4. *Multiple Sclerosis Lesions Recognition: One Class Learning Approach*

Participants: Yogesh Karpate, Olivier Commowick, Christian Barillot, Gilles Edan.

We have developed an automatic algorithm for the detection of multiple sclerosis lesions (MSL) from multi-sequence magnetic resonance imaging (MRI). We build a probabilistic classifier that can recognize MSL as a novel class, trained only on Normal Appearing Brain Tissues (NABT). Patch based intensity information of MRI images is used to train a classifier at the voxel level. The classifier is in turn used to compute a probability characterizing the likelihood of each voxel to be a lesion. This probability is then used to identify a lesion voxel based on simple Otsu thresholding. This work has been submitted to ISBI 2015.

6.5. Vascular Imaging and Arterial Spin Labeling

6.5.1. Peripheral angiography and neurovascular imaging

Participants: Hélène Raoult, Jean-Yves Gauvrit, Elise Bannier, Pierre Maurel, Christian Barillot, Jean-Christophe Ferré.

Work-in-progress Non contrast enhanced MR angiography sequences were optimized on phantom as well as healthy volunteers and evaluated on patients presenting arterio venous malformations (AVM). High temporal resolution (70ms) images were obtained and compared to the gold standard Digital Subtraction Angiography. Results showed that Time-resolved SL MR angiographic imaging over two cardiac cycles is a reliable clinical tool for cerebral AVM characterization, yielding very good to excellent agreement with DSA. This work was published in Radioloy late 2013. This data was also post processed to obtain hemodynamics maps (time to peak, wash-in, wash out and mean transit time) and discriminate among different AVM components to relate hemodynamic patterns with rupture risk. This work was published in Stroke [23].

6.5.2. Robust perfusion maps in Arterial Spin Labeling by means of M-estimators

Participants: Pierre Maurel, Jean-Christophe Ferré, Christian Barillot.

Non-invasive measurement of Cerebral Blood Flow (CBF) is now feasible thanks to the introduction of Arterial Spin Labeling (ASL) Magnetic Resonance Imaging (MRI) techniques. To date, due to the low signal-to-noise ratio of ASL, a single acquisition (pair of control/label scans) is not sufficient to estimate perfusion reliably. Instead, the acquisition is usually repeated several times and the perfusion information is calculated by averaging across the repetitions. However, due to its zero breakdown point, the sample mean is very sensitive to outliers. We have proposed [18] to compute ASL CBF maps using Huber's M-estimator, a robust statistical function that is not overly impacted by outlier. This work was part of the PhD thesis of Camille Maumet.

6.5.3. Brain perfusion gender difference study using MRI in young adults

Participants: Léa Itmi, Pierre Maurel, Christian Barillot.

The usage of population models is becoming increasingly important in cerebral imaging, particularly in ASL. Therefore, it is important to check the limits of the models before applying them, to guarantee the reliability of the results. It is now well-known that brain perfusion changes with the age, and this effect is taken into account when evaluating brain perfusion images. But gender differences have not been well studied yet. It is known for a long time that female brain perfusion is higher than male brain perfusion, but few studies have investigated whether some regional perfusion differences exist or not. We evaluate whether, as for the age, gender differences should be taken into account when analyzing brain perfusion images. We focus on young adults subjects and studied, at the region level and the voxel level if gender differences exist and how it differs. The overall and regional differences were analyzed and then we also investigated the perfusion asymmetries in the brain (left hemisphere versus right hemisphere).

6.6. EEG and MR Imaging

6.6.1. Feasibility and specificity of simultaneous EEG and ASL MRI at 3T

Participants: Elise Bannier, Marsel Mano, Isabelle Corouge, Lorraine Perronnet, Christian Barillot.

Brain functional imaging can be performed using several approaches, including EEG, BOLD and ASL MRI. The Neurinfo platform has acquired an EEG MR compatible 64ch device over the summer to perform joint EEG and BOLD or ASL fMRI. To date, only a few studies have addressed the issue of connecting EEG signal to ASL perfusion. The aim of this study was to assess ASL-EEG at 3T in terms of safety as well as EEG and MR signal quality. The temperature measurements, specific absorption rate, and signal to noise ratio experiments have shown that ASL EEG can be safely performed using the parameters presented above. However, residual gradient artifacts in the PASL-EEG data have to be considered. Further research is needed to understand the artifact variability and to develop an appropriate correction strategy. This study is performed as part of the HEMISFER project in close collaboration with the involved teams.

6.6.2. Neurofeedback using Virtual Reality and Hybrid EEG-MRI for Brain Rehabilitation

Participants: Lorraine Perronnet, Marsel Mano, Christian Barillot.

We have conducted a thorough state-of-the-art of Neurofeedback (NF) and restorative Brain Computer Interfaces (BCI) under EEG and fMRI modality as well as of EEG-fMRI integration, with a particular focus on applications in depression and motor rehabilitation. This enabled us to build a theoretical comparison of EEG- and fMRI-NF methodology that will be helpful in designing NF protocols combining both modalities. In this perspective, we are currently designing a NF protocol based on motor imagery that will be compatible with EEG and fMRI, and running preliminary recordings of motor execution and motor imagery. Besides, we are writing a book chapter about NF and BCI that is intended to disambiguate the existing definitions and to present basic knowledge about NF principles and applications to naive readers. This is a joint work with Anatole Lécuyer team (Hybrid), in the frame of the Hemisfer project.

6.6.3. Symmetrical EEG and fMRI Imaging by Sparse Regularization

Participants: Thomas Oberlin, Pierre Maurel, Christian Barillot.

This work considers the problem of brain imaging using simultaneously recorded electroencephalography (EEG) and functional magnetic resonance imaging (fMRI). To this end, we introduce a linear coupling model that links the electrical EEG signal to the hemodynamic response from the blood-oxygen level dependent (BOLD) signal. Both modalities are then symmetrically integrated, to achieve a high resolution in time and space while allowing some robustness against potential decoupling of the BOLD effect. The novelty of the approach consists in expressing the joint imaging problem as a linear inverse problem, which is addressed using sparse regularization. The sparsity prior naturally reflects the fact that only few areas of the brain are activated at a certain time, and it is easily implemented through efficient so-called proximal algorithms. The significance of the method and the effectiveness of the algorithms are demonstrated through numerical investigations on a simplified head model and simulated data on a realistic brain model. This is a joint work with Remi Gribonval team (Panama), in the frame of the Hemisfer project.

WHISPER Team

6. New Results

6.1. Highlights of the Year

The paper "Faults in Linux 2.6" was published in the ACM journal Transactions on Computer Systems in June 2014 . It has been downloaded from the ACM digital library almost 300 times since then. The paper was reviewed in the Linux Weekly News, in the German professional IT website golem.de, and was the subject of an invited presentation at a joint session of the Linux Kernel Summit and LinuxCon North America.

Julia Lawall was invited to the 2014 Linux Kernel Summit, an invitation-only meeting of core Linux developers. She was subsequently invited to participate in the plenary Linux Kernel Developer Panel at LinuxCon Europe, with 2000 attendees.

Julia Lawall was invited to give a keynote at the conference Modularity (formerly AOSD) on her work on Coccinelle [16].
BEST PAPERS AWARDS :
[] **ACM Transactions on Computer Systems**. N. PALIX, G. THOMAS, S. SAHA, C. CALVÈS, G. MULLER, J. L. LAWALL.

6.2. Lock profiling in Java servers

Today, Java is regularly used to implement large multi-threaded server-class applications that use locks to protect access to shared data. However, understanding the impact of locks on the performance of a system is complex, and thus the use of locks can impede the progress of threads on configurations that were not anticipated by the developer, during specific phases of the execution. In our paper, "Continuously Measuring Critical Section Pressure with the Free-Lunch Profiler" [25], presented at OOPSLA 2014, we propose Free Lunch, a new lock profiler for Java application servers, specifically designed to identify, *in-vivo*, phases where the progress of the threads is impeded by a lock. Free Lunch is designed around a new metric, *critical section pressure* (CSP), which directly correlates the progress of the threads to each of the locks. Using Free Lunch, we have identified phases of high CSP, which were hidden with other lock profilers, in the distributed Cassandra NoSQL database and in several applications from the DaCapo 9.12, the SPECjvm2008 and the SPECjbb2005 benchmark suites. Our evaluation of Free Lunch shows that its overhead is never greater than 6%, making it suitable for *in-vivo* use.

6.3. Software engineering for infrastructure software

A kernel oops is an error report that logs the status of the Linux kernel at the time of a crash. Such a report can provide valuable first-hand information for a Linux kernel maintainer to conduct postmortem debugging. Recently, a repository has been created that systematically collects kernel oopses from Linux users. However, debugging based on only the information in a kernel oops is difficult. In a paper published at MSR [18], we consider the initial problem of finding the offending line, i.e., the line of source code that incurs the crash. For this, we propose a novel algorithm based on approximate sequence matching, as used in bioinformatics, to automatically pinpoint the offending line based on information about nearby machine-code instructions, as found in a kernel oops. Our algorithm achieves 92% accuracy compared to 26% for the traditional approach of using only the oops instruction pointer.

2014 was the second year of a two-year cooperation between Julia Lawall and David Lo of Singapore Management University, as part of the Merlion cooperation grant program of the Insitut Français. This cooperation resulted in four papers: two on word similarity [21], [26], one on bug localization [23], and one on an empirical study of testing practices in open source software [19]. As an offshoot of this work, Julia Lawall worked with the PhD student Ripon Saha of UT Austin and his advisors on the topic of assessing the effectiveness of a state-of-the-art bug localization technique on C programs as compared to Java programs [20]. This work built on the C parser developed for Coccinelle.

Finally, with colleagues from Aalborg University and with Nicolas Palix of Grenoble, Julia Lawall published an article in Science of Computer Programming assessing the applicability of Coccinelle to checking the coding style guidelines of the CERT C Secure Coding Standard [14].

6.4. Bugs in Linux 2.6

In August 2011, Linux entered its third decade. Ten years before, Chou et al. published a study of faults found by applying a static analyzer to Linux versions 1.0 through 2.4.1. A major result of their work was that the drivers directory contained up to 7 times more of certain kinds of faults than other directories. This result inspired numerous efforts on improving the reliability of driver code. Today, Linux is used in a wider range of environments, provides a wider range of services, and has adopted a new development and release model. What has been the impact of these changes on code quality? To answer this question, in an article published in ACM TOCS, we have transported Chou et al.'s experiments to all versions of Linux 2.6; released between 2003 and 2011. We find that Linux has more than doubled in size during this period, but the number of faults per line of code has been decreasing. Moreover, the fault rate of drivers is now below that of other directories, such as arch. These results can guide further development and research efforts for the decade to come. To allow updating these results as Linux evolves, we define our experimental protocol and make our checkers available.

6.5. Memory Monitoring in Smart Home gateways

Smart Home market players aim to deploy component-based and service-oriented applications from untrusted third party providers on a single OSGi execution environment. This creates the risk of resource abuse by buggy and malicious applications, which raises the need for resource monitoring mechanisms. Existing resource monitoring solutions either are too intrusive or fail to identify the relevant resource consumer in numerous multi-tenant situations. In our paper “Memory Monitoring in a Multi-tenant OSGi Execution Environment” [15], presented at CBSE 2014, we propose a system to monitor the memory consumed by each tenant, while allowing them to continue communicating directly to render services. We propose a solution based on a list of configurable resource accounting rules between tenants, which is far less intrusive than existing OSGi monitoring systems. We modified an experimental Java Virtual Machine in order to provide the memory monitoring features for the multi-tenant OSGi environment. Our evaluation of the memory monitoring mechanism on the DaCapo benchmarks shows an overhead below 46%. This work has been done as part of the PhD of Koutheir Attouchi [10] who was supported by a CIFRE grant with Orange Labs.

WILLOW Project-Team

6. New Results

6.1. Highlights of the Year

- J. Sivic started ERC project LEAP (2014-2018).
- J. Sivic serves as a Program Chair for International Conference on Computer Vision, Santiago, Chile, 2015

6.2. 3D object and scene modeling, analysis, and retrieval

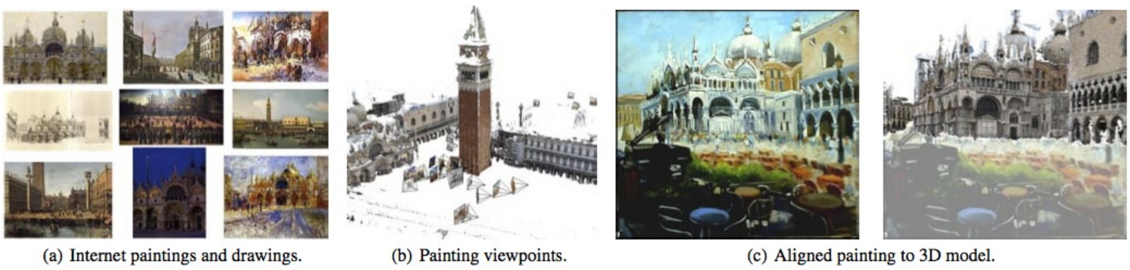


Figure 1. Our system automatically aligns and recovers the viewpoint of paintings, drawings, and historical photographs to a 3D model of an architectural site.

6.2.1. Painting-to-3D Model Alignment Via Discriminative Visual Elements

Participants: Mathieu Aubry, Bryan Russell [Intel Labs], Josef Sivic.

In this work we describe a technique that can reliably align arbitrary 2D depictions of an architectural site, including drawings, paintings and historical photographs, with a 3D model of the site. This is a tremendously difficult task as the appearance and scene structure in the 2D depictions can be very different from the appearance and geometry of the 3D model, e.g., due to the specific rendering style, drawing error, age, lighting or change of seasons. In addition, we face a hard search problem: the number of possible alignments of the painting to a large 3D model, such as a partial reconstruction of a city, is huge. To address these issues, we develop a new compact representation of complex 3D scenes. The 3D model of the scene is represented by a small set of discriminative visual elements that are automatically learnt from rendered views. Similar to object detection, the set of visual elements, as well as the weights of individual features for each element, are learnt in a discriminative fashion. We show that the learnt visual elements are reliably matched in 2D depictions of the scene despite large variations in rendering style (e.g. watercolor, sketch, historical photograph) and structural changes (e.g. missing scene parts, large occluders) of the scene. We demonstrate an application of the proposed approach to automatic re-photography to find an approximate viewpoint of historical paintings and photographs with respect to a 3D model of the site. The proposed alignment procedure is validated via a human user study on a new database of paintings and sketches spanning several sites. The results demonstrate that our algorithm produces significantly better alignments than several baseline methods. This work has been published at ACM Transactions on Graphics 2014 [3] and its extension has appeared at RFIA 2014 [17]. The problem addressed in this work is illustrated in Figure 1 and example results are shown in Figure 2 .

Figure 2. Example alignments of non-photographic depictions to 3D models. Notice that we are able to align depictions rendered in different styles and having a variety of viewpoints with respect to the 3D models.

6.2.2. Seeing 3D chairs: exemplar part-based 2D-3D alignment using a large dataset of CAD models

Participants: Mathieu Aubry, Bryan Russell [Intel labs], Alyosha Efros [UC Berkeley], Josef Sivic.

This work poses object category detection in images as a type of 2D-to-3D alignment problem, utilizing the large quantities of 3D CAD models that have been made publicly available online. Using the "chair" class as a running example, we propose an exemplar-based 3D category representation, which can explicitly model chairs of different styles as well as the large variation in viewpoint. We develop an approach to establish part-based correspondences between 3D CAD models and real photographs. This is achieved by (i) representing each 3D model using a set of view-dependent mid-level visual elements learned from synthesized views in a discriminative fashion, (ii) carefully calibrating the individual element detectors on a common dataset of negative images, and (iii) matching visual elements to the test image allowing for small mutual deformations but preserving the viewpoint and style constraints. We demonstrate the ability of our system to align 3D models with 2D objects in the challenging PASCAL VOC images, which depict a wide variety of chairs in complex scenes. This work has been published at CVPR 2014 [9].

6.2.3. Anisotropic Laplace-Beltrami Operators for Shape Analysis

Participants: Mathieu Andreux [TUM], Emanuele Rodola [TUM], Mathieu Aubry, Daniel Cremers [TUM].

This work introduces an anisotropic Laplace-Beltrami operator for shape analysis. While keeping useful properties of the standard Laplace-Beltrami operator, it introduces variability in the directions of principal curvature, giving rise to a more intuitive and semantically meaningful diffusion process. Although the benefits of anisotropic diffusion have already been noted in the area of mesh processing (e.g. surface regularization), focusing on the Laplacian itself, rather than on the diffusion process it induces, opens the possibility to effectively replace the omnipresent Laplace-Beltrami operator in many shape analysis methods. After providing a mathematical formulation and analysis of this new operator, we derive a practical implementation on discrete meshes. Further, we demonstrate the effectiveness of our new operator when employed in conjunction with different methods for shape segmentation and matching. This work has been published at the Sixth Workshop on Non-Rigid Shape Analysis and Deformable Image Alignment (NORDIA) 2014 [8].

6.2.4. Trinocular Geometry Revisited

Participants: Jean Ponce, Martial Hebert [CMU].

When do the visual rays associated with triplets of point correspondences converge, that is, intersect in a common point? Classical models of trinocular geometry based on the fundamental matrices and trifocal tensor associated with the corresponding cameras only provide partial answers to this fundamental question, in large part because of underlying, but seldom explicit, general configuration assumptions. In this project, we use elementary tools from projective line geometry to provide necessary and sufficient geometric and analytical conditions for convergence in terms of transversals to triplets of visual rays, without any such assumptions. In turn, this yields a novel and simple minimal parameterization of trinocular geometry for cameras with non-collinear or collinear pinholes. This work has been published at CVPR 2014 [15].

6.2.5. On Image Contours of Projective Shapes

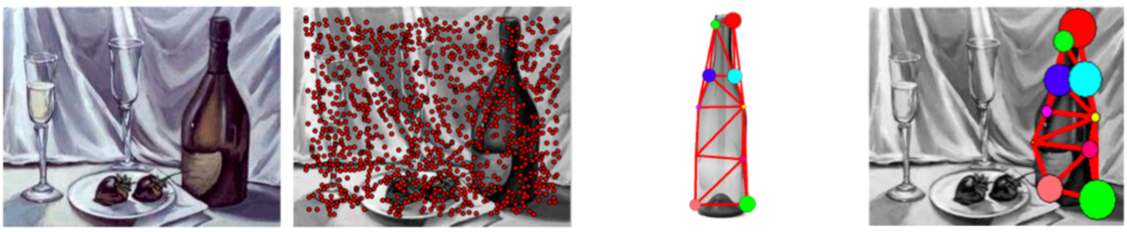
Participants: Jean Ponce, Martial Hebert [CMU].

This work revisits classical properties of the outlines of solid shapes bounded by smooth surfaces, and shows that they can be established in a purely projective setting, without appealing to Euclidean measurements such as normals or curvatures. In particular, we give new synthetic proofs of Koenderink's famous theorem on convexities and concavities of the image contour, and of the fact that the rim turns in the same direction as the viewpoint in the tangent plane at a convex point, and in the opposite direction at a hyperbolic point. This suggests that projective geometry should not be viewed merely as an analytical device for linearizing calculations (its main role in structure from motion), but as the proper framework for studying the relation between solid shape and its perspective projections. Unlike previous work in this area, the proposed approach does not require an oriented setting, nor does it rely on any choice of coordinate system or analytical considerations. This work has been published at ECCV 2014 [14].

6.3. Category-level object and scene recognition

6.3.1. Finding Matches in a Haystack: A Max-Pooling Strategy for Graph Matching in the Presence of Outliers

Participants: Minsu Cho, Jian Sun, Olivier Duchenne, Jean Ponce.



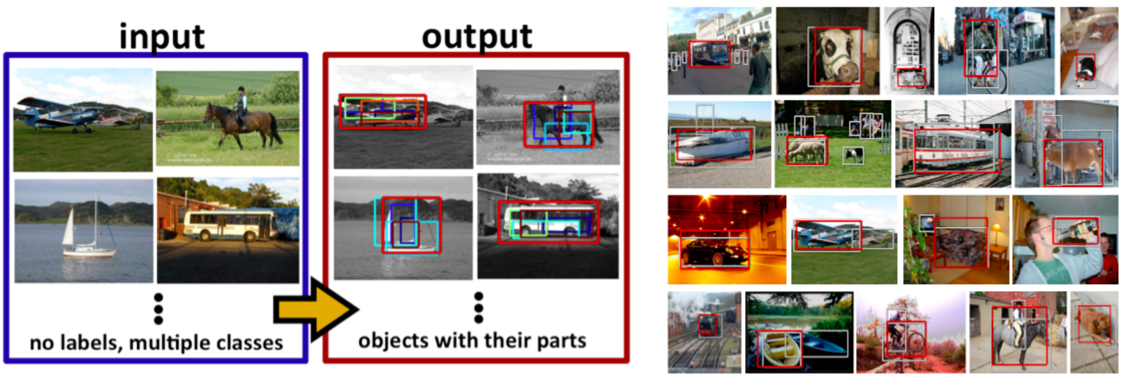
(a) A cluttered scene and its extracted features (b) Feature matching between two images

Figure 3. Feature matching in the presence of outliers. (a) In real-world scenes, background clutter often produces numerous outlier features, making it hard to find correspondences. (b) We address the issue with a max-pooling approach to graph matching. The proposed method is not only resilient to deformations but also remarkably tolerant to outliers. Each node on the left image corresponds to one with the same color on the right image, where bigger nodes represent more similar nodes. (Best viewed in color.)

A major challenge in real-world feature matching problems is to tolerate the numerous outliers arising in typical visual tasks. Variations in object appearance, shape, and structure within the same object class make it harder to distinguish inliers from outliers due to clutters. In this work, we propose a max-pooling approach to graph matching, which is not only resilient to deformations but also remarkably tolerant to outliers. The proposed algorithm evaluates each candidate match using its most promising neighbors, and gradually propagates the corresponding scores to update the neighbors. As final output, it assigns a reliable score to each match together with its supporting neighbors, thus providing contextual information for further verification. We demonstrate the robustness and utility of our method with synthetic and real image experiments. This work has been published at CVPR 2014 [11]. The proposed method and its qualitative results are illustrated in Figure 3 .

6.3.2. Unsupervised Object Discovery and Localization in the Wild: Part-based Matching with Bottom-up Region Proposals

Participants: Minsu Cho, Suha Kwak, Cordelia Schmid [Inria Lear], Jean Ponce.



(a) Fully unsupervised object localization (b) Localization examples from PASCAL VOC 2007

Figure 4. Unsupervised object discovery in the wild. (a) We tackle object localization in an unsupervised scenario without any types of annotations, where a given image collection may contain multiple dominant object classes and even outlier images. The proposed method discovers object instances (red bounding boxes) with their distinctive parts (smaller boxes). (b) Examples of localization on mixed-class PASCAL VOC 2007 train/val datasets are shown. Red boxes represent localized objects while white boxes are ground truth annotations. (Best viewed in color.)

This work addresses unsupervised discovery and localization of dominant objects from a noisy image collection of multiple object classes. The setting of this problem is fully unsupervised, without even image-level annotations or any assumption of a single dominant class. This is significantly more general than typical colocalization, cosegmentation, or weakly-supervised localization tasks. We tackle the discovery and localization problem using a part-based matching approach: We use off-the-shelf region proposals to form a set of candidate bounding boxes for objects and object parts. These regions are efficiently matched across images using a probabilistic Hough transform that evaluates the confidence in each candidate region considering both appearance similarity and spatial consistency. Dominant objects are discovered and localized by comparing the scores of candidate regions and selecting those that stand out over other regions containing them. Extensive experimental evaluations on standard benchmarks demonstrate that the proposed approach significantly outperforms the current state of the art in colocalization, and achieves robust object discovery in challenging mixed-class datasets. This work has been submitted to CVPR 2015 [22]. The proposed method and its qualitative results are illustrated in Figure 4 .

6.3.3. Learning and Transferring Mid-Level Image Representations using Convolutional Neural Networks

Participants: Maxime Oquab, Leon Bottou [MSR New York], Ivan Laptev, Josef Sivic.

Convolutional neural networks (CNN) have recently shown outstanding image classification performance in the large-scale visual recognition challenge (ILSVRC2012). The success of CNNs is attributed to their ability to learn rich mid-level image representations as opposed to hand-designed low-level features used in other image classification methods. Learning CNNs, however, amounts to estimating millions of parameters and requires a very large number of annotated image samples. This property currently prevents application of CNNs to problems with limited training data. In this work we show how image representations learned with CNNs on large-scale annotated datasets can be efficiently transferred to other visual recognition tasks with limited amount of training data. We design a method to reuse layers trained on the ImageNet dataset to compute mid-level image representation for images in the PASCAL VOC dataset. We show that despite differences in image statistics and tasks in the two datasets, the transferred representation leads to significantly improved results for object and action classification, outperforming the current state of the art on Pascal VOC 2007 and 2012 datasets. We also show promising results for object and action localization. This work has been published at CVPR 2014 [13].

6.3.4. Weakly supervised object recognition with convolutional neural networks

Participants: Maxime Oquab, Leon Bottou [MSR New York], Ivan Laptev, Josef Sivic.

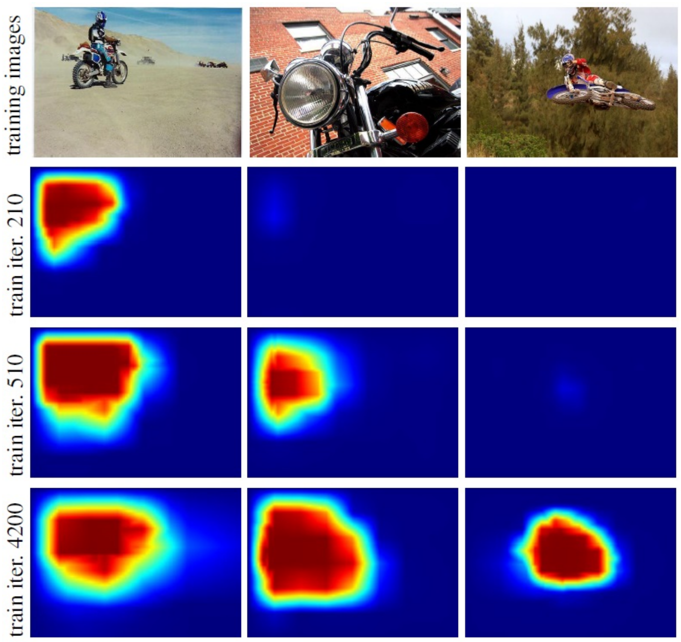


Figure 5. Evolution of localization score maps for the motorbike class over iterations of our weakly-supervised CNN training. Note that locations of objects with more usual appearance are discovered earlier during training.

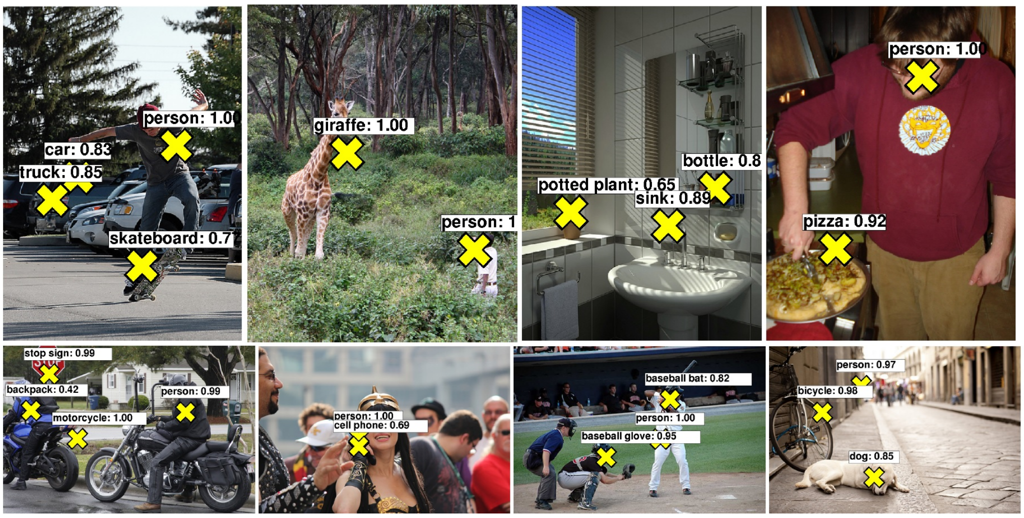


Figure 6. Example location predictions for images from the Microsoft COCO validation set obtained by our weakly-supervised method. Note that our method does not use object locations at training time, yet can predict locations of objects in test images (yellow crosses). The method outputs the most confident location for most confident object classes.

Successful methods for visual object recognition typically rely on training datasets containing lots of richly annotated images. Detailed image annotation, e.g. by object bounding boxes, however, is both expensive and often subjective. We describe a weakly supervised convolutional neural network (CNN) for object classification that relies only on image-level labels, yet can learn from cluttered scenes containing multiple objects (see Figure 5). We quantify its object classification and object location prediction performance on the Pascal VOC 2012 (20 object classes) and the much larger Microsoft COCO (80 object classes) datasets. We find that the network (i) outputs accurate image-level labels, (ii) predicts approximate locations (but not extents) of objects, and (iii) performs comparably to its fully-supervised counterparts using object bounding box annotation for training. This work has been submitted to CVPR 2015 [23]. Illustration of localization results by our method in Microsoft COCO dataset is illustrated in Figure 6 .

6.3.5. Learning Dictionary of Discriminative Part Detectors for Image Categorization and Cosegmentation

Participants: Jian Sun, Jean Ponce.

This work proposes a novel approach to learning mid-level image models for image categorization and cosegmentation. We represent each image class by a dictionary of discriminative part detectors that best discriminate that class from the background. We learn category-specific part detectors in a weakly supervised setting in which the training images are only labeled with category labels without part / object location labels. We use a latent SVM model regularized by l1,2 group sparsity to learn the discriminative part detectors. Starting from a large set of initial parts, the group sparsity regularizer forces the model to jointly select and optimize a set of discriminative part detectors in a max-margin framework. We propose a stochastic version of a proximal algorithm to solve the corresponding optimization problem. We apply the learned part detectors to image classification and cosegmentation, and quantitative experiments with standard benchmarks show that our approach matches or improves upon the state of the art. This work has been submitted to PAMI [24].

6.4. Image restoration, manipulation and enhancement

6.4.1. Fast Local Laplacian Filters: Theory and Applications

Participants: Mathieu Aubry, Sylvain Paris [Adobe], Samuel Hasinoff [Google], Jan Kautz [University College London], Fredo Durand [MIT].

Multi-scale manipulations are central to image editing but they are also prone to halos. Achieving artifact-free results requires sophisticated edge-aware techniques and careful parameter tuning. These shortcomings were recently addressed by the local Laplacian filters, which can achieve a broad range of effects using standard Laplacian pyramids. However, these filters are slow to evaluate and their relationship to other approaches is unclear. In this work, we show that they are closely related to anisotropic diffusion and to bilateral filtering. Our study also leads to a variant of the bilateral filter that produces cleaner edges while retaining its speed. Building upon this result, we describe an acceleration scheme for local Laplacian filters on gray-scale images that yields speed-ups on the order of 50x. Finally, we demonstrate how to use local Laplacian filters to alter the distribution of gradients in an image. We illustrate this property with a robust algorithm for photographic style transfer. This work has been published at ACM Transactions on Graphics 2014 [2].

6.4.2. Learning a Convolutional Neural Network for Non-uniform Motion Blur Removal

Participants: Jian Sun, Wenfei Cao, Zongben Xu, Jean Ponce.

In work work, we address the problem of estimating and removing non-uniform motion blur from a single blurry image. We propose a deep learning approach to predicting the probabilistic distribution of motion blur at the patch level using a convolutional neural network (CNN). We further extend the candidate set of motion kernels predicted by the CNN using carefully designed image rotations. A Markov random field model is then used to infer a dense non-uniform motion blur field enforcing the motion smoothness. Finally the motion blur is removed by a non-uniform deblurring model using patch-level image prior. Experimental evaluations show that our approach can effectively estimate and remove complex non-uniform motion blur that cannot be well achieved by the previous approaches. This work has been submitted to CVPR 2015.

6.5. Human activity capture and classification

6.5.1. Weakly Supervised Action Labeling in Videos Under Ordering Constraints

Participants: Piotr Bojanowski, Remi Lajugie [Inria Sierra], Francis Bach [Inria Sierra], Ivan Laptev, Jean Ponce, Cordelia Schmid [Inria Lear], Josef Sivic.

We are given a set of video clips, each one annotated with an ordered list of actions, such as “walk” then “sit” then “answer phone” extracted from, for example, the associated text script. We seek to temporally localize the individual actions in each clip as well as to learn a discriminative classifier for each action. We formulate the problem as a weakly supervised temporal assignment with ordering constraints. Each video clip is divided into small time intervals and each time interval of each video clip is assigned one action label, while respecting the order in which the action labels appear in the given annotations. We show that the action label assignment can be determined together with learning a classifier for each action in a discriminative manner. We evaluate the proposed model on a new and challenging dataset of 937 video clips with a total of 787720 frames containing sequences of 16 different actions from 69 Hollywood movies. This work has been published at ECCV 2014 [10].

6.5.2. Predicting Actions from Static Scenes

Participants: Tuan-Hung Vu, Catherine Olsson [MIT], Ivan Laptev, Aude Oliva [MIT], Josef Sivic.

Human actions naturally co-occur with scenes. In this work we aim to discover action-scene correlation for a large number of scene categories and to use such correlation for action prediction. Towards this goal, we collect a new SUN Action dataset with manual annotations of typical human actions for 397 scenes. We next discover action-scene associations and demonstrate that scene categories can be well identified from their associated actions. Using discovered associations, we address a new task of predicting human actions for images of static scenes. We evaluate prediction of 23 and 38 action classes for images of indoor and outdoor scenes respectively and show promising results, see Figure 7 . We also propose a new application of geo-localized action prediction and demonstrate ability of our method to automatically answer queries such as “Where is a good place for a picnic?” or “Can I cycle along this path?”. This work has been published in ECCV 2014 [16].

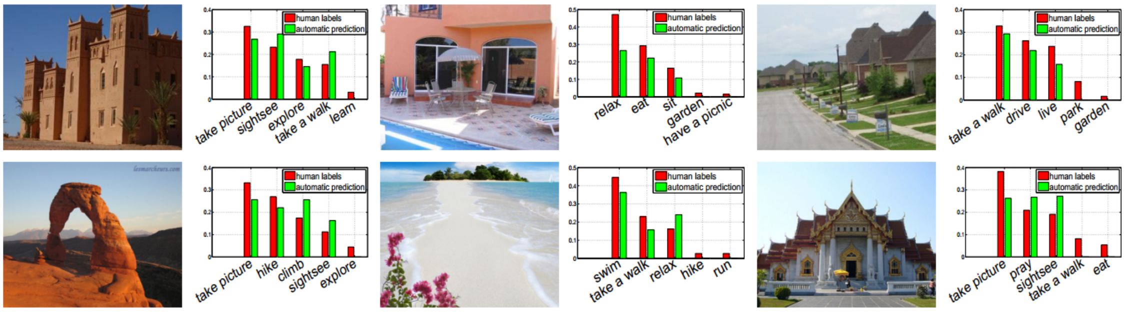


Figure 7. Automatic visual action prediction for test images in SUN Action dataset.

6.5.3. Efficient feature extraction, encoding and classification for action recognition

Participants: Vadim Kantorov, Ivan Laptev.

Local video features provide state-of-the-art performance for action recognition. While the accuracy of action recognition has been continuously improved over the recent years, the low speed of feature extraction and subsequent recognition prevents current methods from scaling up to real-size problems. We address this issue and first develop highly efficient video features using motion information in video compression. We next explore feature encoding by Fisher vectors and demonstrate accurate action recognition using fast linear classifiers. Our method improves the speed of video feature extraction, feature encoding and action classification by two orders of magnitude at the cost of minor reduction in recognition accuracy. We validate our approach and compare it to the state of the art on four recent action recognition datasets. This work has been published at CVPR 2014 [12].

6.5.4. On Pairwise Cost for Multi-Object Network Flow Tracking

Participants: Visesh Chari, Simon Lacoste-Julien [Inria Sierra], Ivan Laptev, Josef Sivic.

Multi-object tracking has been recently approached with the min-cost network flow optimization techniques. Such methods simultaneously resolve multiple object tracks in a video and enable modeling of dependencies among tracks. Min-cost network flow methods also fit well within the "tracking-by-detection" paradigm where object trajectories are obtained by connecting per-frame outputs of an object detector. Object detectors, however, often fail due to occlusions and clutter in the video. To cope with such situations, we propose an approach that regularizes the tracker by adding second order costs to the min-cost network flow framework. While solving such a problem with integer variables is NP-hard, we present a convex relaxation with an efficient rounding heuristic which empirically gives certificates of small suboptimality. Results are shown on real world video sequences and demonstrate that the new constraints help selecting longer and more accurate tracks improving over the baseline tracking-by-detection method. This work has been submitted to CVPR 2015 [21].

WIMMICS Project-Team

6. New Results

6.1. Highlights of the Year

BEST PAPERS AWARDS :

[52] **IEEE/WIC/ACM International Conference on Intelligent Agent Technology, IAT**. A. G. B. TETTAMANZI, C. DA COSTA PEREIRA.

[45] **10th IEEE International Conference Beyond Databases, Architectures, and Structures (BDAS 2014)**. T. H. H. NGUYEN, N. LE THANH.

6.2. Argumentation Theory

6.2.1. Combining Argumentation Theory and Normative Reasoning with Natural Language Processing

Participants: Serena Villata, Elena Cabrio, Fabien Gandon.

We have proposed a methodology to identify and classify the semantic relations holding among the possible different answers obtained for a certain query on DBpedia language specific chapters. The goal is to reconcile information provided by language specific DBpedia chapters to obtain a consistent results set. The results of this research have been published at the LREC conference [29]. This classification has then been exploited in another work, together with Elena Cabrio and Alessio Palmero Aprosio (FBK Trento, Italy), where Serena Villata has worked on an extension of QAKiS, the system for open domain Question Answering over linked data, that allows to query DBpedia multilingual chapters. Such chapters can contain different information with respect to the English version, e.g. they provide more specificity on certain topics, or fill information gaps. In particular, she extended the results presented last year embedding the new identified relations among the different answers, using argumentation theory to reconcile the information and further improving the system's performances. A demo of the new argumentation module is available online [0]. This work has also been presented at the International Semantic Web Conference demo session [85].

Moreover, we have proposed, together with Alessio Palmero Aprosio, a system called NLL2RDF to translate in an automated way licenses, such as GPL, in natural language into a machine-readable version using the RDF language. The system is available online [0]. The results of this research have been presented at the European Semantic Web conference [26].

Finally, we have published the benchmark of natural language arguments called NoDE. The benchmark is available online [0]. The results of this research have been presented at the 15th International Workshop on Non-Monotonic Reasoning (NMR 2014) [28] and at the 5th Conference on Computational Argumentation [27] (COMMA 2014 - demo).

6.2.2. Argumentation and Legal Reasoning

Participant: Serena Villata.

[0] http://qakis.org/qakis2/

[0] http://www.airpedia.org/nll2rdf/

[0] http://www-sop.inria.fr/NoDE/

Together with Leendert van der Torre (University of Luxembourg), we proposed a framework for reasoning about norms using argumentation theory. Norms regulate our everyday life, and are used to assess the conformance of our behavior with respect to the regulations holding in specific contexts. Given the profound importance of norms in our lives, it is fundamental to understand which norms are valid in certain environments, how to interpret them, the legal conclusions of such norms, which norms can be derived from the existing ones, etc. In order to understand norms, people discuss about them to assess the validity or applicability of a certain norm subject to particular conditions, to derive the obligations and permissions to be enforced, or claim that a certain normative conclusion cannot be derived from the existing regulations. Several frameworks have been proposed for legal argumentation, but no comprehensive formal model of legal reasoning from arguments has been proposed yet. The goal of this work is to enrich legal argumentation with a formal account of deontic modalities. These results have been published at the 5th Conference on Computational Argumentation [55] (COMMA 2014).

Moreover, together with Guido Boella (University of Torino, Italy), Pietro Baroni and Massimiliano Giacomin (University of Brescia, Italy), Federico Cerutti (University of Aberdeen, UK), Leendert van der Torre (University of Luxembourg), we have studied also the dynamics of argumentation framework and this research has lead to a publication in the Artificial Intelligence journal [15].

6.3. License for the Web of Data

6.3.1. Reasoning about rights and licenses in the Web of Data

Participant: Serena Villata.

In the domain of Linked Open Data a need is emerging for developing automated frameworks able to generate the licensing terms associated to data coming from heterogeneous distributed sources. Together with Guido Governatori (NICTA, Australia) and Antonino Rotolo (University of Bologna, Italy), Serena Villata proposed and evaluated a deontic logic semantics which allows to define the deontic components of the licenses, i.e., permissions, obligations, and prohibitions, and generate a composite license compliant with the licensing items of the composed different licenses. The approach is evaluated using the SPINdle defeasible reasoner, where the proposed heuristics have been hard coded in the reasoner. The prosecution of this research line has seen the analysis of the compatibility of a set of licensing terms (always using SPINdle), the analysis of the role of licenses associated to vocabularies, and the development of the Licentia suite of services to reason over licenses and help users to deal with such kind of information. The results of this research line have been published at the International Semantic Web Conference demo session [87].

6.4. Logic and Cognition

6.4.1. Logical Foundations of Cognitive Agents

Participants: Andrea Tettamanzi, Serena Villata.

Together with Célia da Costa Pereira of I3S, we have investigated syntactic belief revision operators [16] and goal-generation mechanisms [30] to make the practical implementation of a general BDI (Belief-Desire-Intention) model of agency based on possibility theory. Furthermore, we took part in a joint investigation with a research team, led by Cristiano Castelfranchi, of the CNR-ISTC in Rome on the issue of trust in multi-agent systems [21]. We also employed agent-based simulation to test a theory of human stupidity proposed by the late Italian economist Carlo Cipolla [52]; our paper won the Best Paper Award at IAT 2014.

6.4.2. RDF Mining

Participants: Andrea Tettamanzi, Catherine Faron Zucker, Fabien Gandon.

We carried on our investigation in an approach to RDF mining based on grammatical evolution and possibility theory, whose aim is to mine large RDF graphs by automatically generating and testing OWL 2 axioms based on the known facts. In particular, we addressed the problem of testing candidate OWL 2 axioms against the fact contained in an RDF base and proposed a novel scoring heuristics based on falsification and possibility theory [53].

6.4.3. Sentiment Analysis

Participant: Andrea Tettamanzi.

Together with Célia da Costa Pereira of I3S and Mauro Dragoni of FBK, Trento, who visited our team for three months from April to June 2014, we have proposed a novel approach to concept-level sentiment analysis based on fuzzy logic. Our system [31], [89] participated in the Semantic Web Evaluation Challenge (SemWebEval) at ESWC 2014 and was the winner for Task 1 and Most Innovative Approach.

6.4.4. Data and Knowledge Integration and Extraction

Participant: Andrea Tettamanzi.

Together with Somsack Inthasone, Nicolas Pasquier and Célia da Costa Pereira of I3S, we developed a data warehouse collecting data for research on biodiversity [38].

6.4.5. Miscellaneous

Participant: Andrea Tettamanzi.

A work on electorcardiographic signal classification using evolutionary algorithms and neural networks carried out while still at the University of Milan, got published as a book chapter [65].

6.5. Natural Language Processing

6.5.1. Bridging Natural Language Processing and the Web of Data for Multimedia Question Answering

Participants: Elena Cabrio, Fabien Gandon, Yoann Moise.

Differently from search engines, the goal of Question Answering (QA) is to return precise answers to users' natural language questions, extracting information from both documentary text and advanced media content. Up to now, QA research has largely focused on text, mainly targeting factual and list questions. The goal of our work was instead to exploit structured data and metadata describing multimedia content on Linked Open Data to provide a richer and more complete answer to the user, combining textual information with other media content.

We implemented an extension of our QAKiS system [0] to boost the answer visualization adding multimedia content. More specifically, once QAKiS outputs the textual answer(s) to the question asked by the user, the user can click on "more details" to have further information on the retrieved entity. Three main types of additional (and multimedia) content are then displayed: *i)* additional textual information providing a description of the retrieved entity (extracted from DBpedia), and a structured Information Card containing a set of relevant properties of the entity; *ii)* images (extracted from Flickr) and relevant videos (extracted from YouTube); *iii)* entity geo-localization: a pointer on a map is shown for questions asking about a place, together with its points of interest. The results of this research have been published at ESWC 2014 - Demo/poster paper [84].

6.5.2. SMILK - Social Media Intelligence and Linked Knowledge

Participants: Elena Cabrio, Fabien Gandon, Fabrice Jauvat.

Automated Natural Language Processing (NLP), Web Open Data (Linked Open Data) and social networks are the three topics of the SMILK ANR LabCom including their coupling studied in three ways: texts and Linked Data, Linked Data and social resources, texts and social resources. The purpose of this LabCom is indeed to develop research and technologies on the one hand, retrieve, analyze, and reason about linking data from textual Web resources and other to use open Web data taking into account the social structures and interactions in order to improve the analysis and understanding of textual resources.

[0] http://qakis.org

As a first step in this direction, during the internship of Fabrice Jauvat we have developed a prototype of a system that - given free text (in particular in the cosmetics domain, extracted from a forum, a magazine, or a Web site) - can first recognize the named entities launching in parallel the RENCO system (developed by our partner in the LabCom), and NERD [0], and then connect them to DBpedia, so that additional information on the entity can be extracted. For instance, if the word "J'adore" is detected in the text, it is recognized as a named entity and connected to its DBpedia page, so that information about the fact that it is a perfume, that its brand is Dior, and so on, can be automatically extracted and shown to the user.

6.5.3. Ontology-Based and Natural Language Chatbot System in the Commercial Domain

Participants: Amine Hallili, Catherine Faron Zucker, Elena Cabrio, Fabien Gandon.

This work is done within a Cifre PhD Thesis colocated in the Wimmics team and with SynchroNext Company located in Nice. The work consists in modelling and implementing ontology-based natural language Chatbot in commercial domain which consists of

- The design of a commercial knowledge base using the websites' APIs and web services (e.g. Amazon1, eBay2, BestBuy3),
- Interpreting and handling links between users' natural language questions by constructing relational graph,
- Generation and visualization of textual and media answers.

A. Hallili attended the ESSLLI summer school where a poster was accepted [93].

6.5.4. Editor of Formal Lexicographic Definitions

Participants: Maxime Lefrançois, Romain Gugert, Alain Giboin, Fabien Gandon.

Last year a prototype of a GUI of an editor of formal dictionary definitions aimed at lexicographers was developed based on the formalism of Units Graphs and on Meaning-Text Theory. This year, the prototype was demonstrated during the IC 2014 conference [60]. The prototype was also described in a paper reporting the knowledge engineering methodology for representing lexicographic definitions it supports [39].

6.6. Collaborative Software Development Platforms

Participant: Isabelle Mirbel.

Today's Web has given rise to several platforms serving the purpose of collaborative software development. Thanks to these environments, it is possible, among others, for anyone to suggest new requirements for a software under development. A lot of requirements are thus proposed by users and it becomes difficult, after a while, for the persons in charge of the software which development is hosted by the platform to understand this large set of new requirements in its entirety. Therefore we proposed a tool to make large sets of requirements posted on collaborative software development platforms better workable despite the poor content of requirement body. Our aim was to propose an approach to automatically group similar requirements together in order to propose a limited number of requirement categories, thus improving the review process. As requirements expressed on collaborative software development platforms are usually very short and their content not very structured, we proposed to exploit relationships between stakeholders and already processed requirements to break the whole set of new requirements into meaningful categories. Our tool relies on Semantic Web languages and Formal Concept Analysis to provide a 3 steps data analysis process. The data is first extracted from the platform and translated into RDF, then stakeholders' past activities are analyzed to finally get stakeholder categories in order to improve the review of newly posted requirements.

According to the experiments that we conducted, we noticed some limitations in our approach. When the contributing stakeholders are newbies with no previous participation in any blueprint or bug and when there is no sufficient number of evaluated blueprints or bugs. To cope with this limitation, we plan to evaluate stakeholders reputation by looking at their activities on the whole collaborative software development platform (and not only the project under consideration). The results of this research have been published in [23].

[0] http://nerd.eurecom.fr/

In today's software development methodologies, User Stories (US) are mostly used as primary requirements artifacts. They are used to express requirements from a final user point of view and at a low abstraction basis using natural language. Over the years, several informal templates have been proposed by agile methods practitioners or academics to guide requirements gathering. Consequently, these templates are used in an ad-hoc manner, each modeler having idiosyncratic preferences. In this context, we performed a study of templates found in literature in order to propose a unified model. We also proposed an RDFS translation of this model in order to allow the annotation of user stories, thus providing search and reasoning capabilities to agile methods practitioners. The results of this research have been published in [56].

6.7. Workflow Management Systems

6.7.1. Semantic Mappings with a control flow-based business workflow

Participants: Thi Hoa Hue Nguyen, Nhan Le Thanh.

The aim of this PhD work is to improve Coloured Petri Nets (CPNs) and Ontology engineering to support the development of business process and business workflow definitions of various fields. To realize this objective, in the first time, we propose an ontological approach for representing business models in a meta-knowledge base. We introduce four basic types of manipulation operations on process models used to develop and modify business workflow patterns. In the second time we propose a formal definition of semantic constraints and an O(n3)-time algorithm for detecting redundant and conflicting constraints. By relying on the CPN Ontology and sets of semantic constraints, workflow processes are semantically created. Finally, we show how to check the semantic correctness of workflow processes with the SPARQL query language [45], [20], [46].

6.7.2. Extraction Mechanisms and Semantic Enrichment of Short Messages in Social Networks

Participants: Amosse Edouard, Nhan Le Thanh.

In this PhD, the work is focused on text processing in social networks and the main objectives are focused on the analysis of the spatial aspect, context enrichment and spatiotemporal analysis of short text messages.

During the first half of the year, we have mainly worked on positioning the research subject beside the state of the art as well as determining relevant domain. After analyzing several works on short text analysis in many domains such as Semantic Web, Data Mining and Natural Language Processing, we have identified a lack in the representation of the spatial aspect. Indeed, the spatial properties of items shared among online communities can be seen on three different aspects: i) The location of the resources which can be identified by its URI/URL, ii) The producer's location, iii) The location related to the content of the messages.

Most existent works have considered as identical the producer's location and the event described by the content of the message, which can lead to wrong results in many cases. For example, a user can be in the United States while describing an event in Africa. The SIOC ontology is one of the most known for representing items shared among online communities; we have proposed an extension of this ontology in which the spatial aspects are clearly represented. However, there exists a big challenge in finding the relevant location that can be associated to the content of a message. We are currently working on an approach that combines NLP technics and GIS to identify the spatial location of an item by analyzing its content.

6.7.3. Ontology-Based Workflow Management Systems

Participants: Tuan Anh Pham, Nhan Le Thanh.

The main objective of the PhD work is to develop a Shared Workflow Management System (SWMS) using ontology engineering. Everybody can share a semi-complete workflow which is called *Workflow template*, and other people can modify and complete it to use in their system. This customized workflow is called *Personalized workflow*. The challenges of a SWMS is to be simple, easy to use, friendly with the user and not too heavy. But it must have all functions of a WMS. There are three major challenge in this work: How to allow the users to customize the workflow template to correspond to their requirements, but their changes must be compliance with the predefined rules in the workflow template? How to build an execution model to evaluate step by step a personalized workflow ?

6.7.4. Model Spatio-Temporal Dedicated Social Networks

Participants: Amel Ben Othmane, Nhan Le Thanh.

The research aims of the work are to: i) model spatio-temporal, dedicated social networks using semantic web models (ontologies) taking into account spatial, temporal, social and dedicated dimensions. ii) overcome limitations of traditional Recommender Systems and improve the quality of recommendation by exploiting context (time, location, goal, etc.) and social ties.

The following tasks, proposed in the first year planning, are completed or almost finished, and are highly relevant to the current work, despite the different initial overall aim:

- Literature Review Report in Social networks, Spatio-temporal networks, Dedicated networks, Activities modeling techniques, Recommender systems.
- Elaboration of the requirements of an ”ideal system” and presentation of an initial approach. The approach named the 5ws approach tries to answer those five questions: who must do, what, when, where and why?
- Implementation of the approach with Protégé.
- Extend semantic sensor network ontology to meet our requirements. In fact we use this ontology to enrich data from sensors networks which are used to measure different metrics describing physical activities (speed, heart rates, distance, etc.).

The following developments are ongoing: Adaptation of recommender systems for activities recommendation and reusing multi-dimensional recommendation model.

6.8. Modeling Team Processes

6.8.1. Modeling and Assessing Coordination Processes

Participants: Alain Giboin, Isabelle Mirbel.

This work is done in collaboration with Pierre Robillard (Polytechnique Montréal).

Last year we worked on an assessment method of the quality of team dynamics based on a taxonomy of episodes of interactions encountered in software development teams [99] – the CoDyMA (Collaborative Dynamics Measurement and Analysis) method. Precisely, we proposed an analysis procedure of episodes based on the Formal Concept Analysis (FCA) approach. This year, we proposed to enrich the CoDyMA method with a procedure for assessing the quality of coordination interactions and the quality of coordination artifacts within a development team. The procedure is based on the ”Coordinative Artifacts” Framework [101], [100].

6.8.2. Modeling Multimodal Grounding Processes in Design Teams

Participant: Alain Giboin.

This work is done with Aurore Defays (Université de Liège).

Grounding is the process used by participants to a collective activity to coordinate both the content and process of their communication to be successful [96]. Grounding is also defined as the process of elaborating and maintaining the Common Ground (i.e., mutual knowledge, mutual beliefs, and mutual assumptions) necessary to participants’ mutual understanding [97]. Multimodal grounding is the process of grounding using several perceptual modalities. Last year we improved the methodology of analysis of multimodal grounding proposed in [98].

6.9. Semantic and Temporal Analysis of Online Communities

Participant: Zide Meng.

The objective of the OCKTOPUS ANR project is to increase the potential social and economic benefit of the large and quickly growing amounts of user-generated content, by transforming it into useful knowledge. Since user communities are the basic of user-generated content sites, we start with community detection problem, which is a fundamental research point in social network analysis. Based on the preliminary experience from the previous year, we made several progress this year and published the results in international conferences, specifically:

- Topic based interested group detection:
  By analyzing a dataset extracted from the popular question answer site "StackOverflow", we proposed a heuristic method to enrich questions' tag. We also introduced a tag tree based model to extract topics from questions' tags, then we used the detected topics to label users in order to detect interest groups. We conducted experiments on the dataset and compared with related method. Results show that the proposed method is much simple and fast. This work has been published in [43].
- Question Answer social media management
  We proposed a question answer social media system based on social network analysis and social media mining to manage the two main resources in question answer sites: users and contents. We also presented a vocabulary used to formalize both the level of interest and the expertise of users on topics. We tested QASM on a dataset extracted from the popular "StackOverflow" site. We showed how the formalized knowledge is used to find relevant experts for a question. This work has been published in [95].

6.9.1. Temporal analysis in User and Topic:

We are planning to introduce temporal analysis into our research problem. According to the previous work, the potential direction could be topic evolution and user interest evolution. We believe this work could benefit community management in question answer sites, for example topic trend detection or user interest management.

6.10. P2P Media Streaming

Participant: Gaspard Perrot.

The *Heave-Ho* project won the Inria 2014 *Boost Your Code* contest.

The goal of the project is to design an overlay network for P2P media streaming based on new HTML5 technologies such as WebRTC. While conventional Internet applications encounter problems with scaling up as the number of visitors grows, the Heave-Ho project aims to enable website's users to share the resource directly among them. The proposed solution is a perfect fit for real-time video broadcasting. In traditional server/client architecture the server can only handle a limited number of requests; if there are too many clients, some of them will not have access to the video. Using a P2P system, the video can be broadcasted to more clients. The use of sharing techniques based on user location can also cut data transfer costs directly at ISP level, thereby reducing the risk of problems such as data rate limits [0].

6.11. Discovery Hub

Participants: Nicolas Marie, Fabien Gandon, Emilie Palagi, Alain Giboin.

In the context of the Discovery Hub project [0], we performed an ergonomic evaluation and redesign of the graphical and textual User Interfaces (UIs) displaying, and allowing to interact with, the explanations provided by Discovery Hub to justify the results it has retrieved. We did user tests of the existing UIs and designed several mockups improving the UIs by taking users' feedback into account.

[0] https://www.inria.fr/en/content/view/full/61889
[0] http://www.discoveryhub.co

We also performed a user-centered evaluation of the quality of the results retrieved by the 4 algorithms of Discovery Hub. We decided to focus on the quality of the results and not on the UI. Thus, specific criteria of the quality of the results were defined in this evaluation: the surpriseness and the interestingness of the results. A result is considered as:

- surprising if the user discovers an unknown resource or relation between the topic searched and the selected result, or if she discovers something unexpected ;

- interesting if the user thinks it is similar to the topic explored or if she thinks she will remind or reuse it later on.

We are currently developing an ergonomic method for evaluating exploratory search systems (ESSs) in general. We are performing a first test of the method on Discovery Hub [41].

6.12. Knowledge Graphs

6.12.1. SPARQL Template Transformation Language

Participants: Olivier Corby, Catherine Faron-Zucker Song.

We finalized the design and implementation of SPARQL Template Transformation Language [0] (STTL) [58], [73]. STTL is an extension of SPARQL with a `template {} where {}` clause that enables a transformation to generate presentation format for RDF. For example, it is possible to generate Turtle, OWL or SPIN-to-SPARQL syntax.

We designed a new service in Corese server that returns HTML. Using STTL transformations that generate HTML, we are able to set up light weight Semantic Web servers on top of local RDF Datasets or remote Datasets such as DBpedia [0].

6.12.2. RDF Serialization and Introspection

Participant: Olivier Corby.

We started a work on RDF serialization of (Java) objects for Semantic Web system introspection. In conjunction with the overloading of SPARQL named graph pattern, we are able to query the system on several internal status such as graph index, triple provenance, property path triples, etc.

6.12.3. OWL 2 RL

Participant: Olivier Corby.

We dramatically optimized Corese Inference Rule Engine and we were able to run OWL 2 RL [0] rule base on the FMA ontology (Fundational Model of Anatomy [0]) with interesting performance. The initial OWL graph contains 1.74 million triples, the final graph contains 13.46 million triples and the rule engine runs in 3 minutes.

6.12.4. Rules for the Web of Data

Participants: Oumy Seye, Olivier Corby, Catherine Faron Zucker, Fabien Gandon.

This year we focused on the validation, the update of rules bases and the optimization of the reasonning. The goal of this work is to detect some inconsistencies in selected rule bases with respect to ontology and offer users to correct this. We built a set of SPARQL queries enabling (1) to build specific rule bases for a given context or application, (2) to optimize inference engines based on rule selection with respect to target RDF data sources, (3) to validate and update rule bases. We propose another optimization of the inference engines based on graph of rules dependencies and rules aplication ordering. This work is published in [62].

[0] http://ns.inria.fr/sparql-template/
[0] http://corese.inria.fr
[0] http://www.w3.org/TR/owl2-profiles/#OWL_2_RL
[0] http://sig.biostr.washington.edu/projects/fma/release/v3.2.1/alt_formats.html

6.12.5. KGRAM

Participants: Olivier Corby, Fuqi Song.

We received a two years grant from Inria to support the development of the Corese platform. This action aims at enhancing Corese software for conforming to latest W3C standards and facilitating its usage in distributed environment, we integrated several open source parsers to Corese, including JSON-LD, RDFa, TriG and N-Quads. Corese now is able to process RDF dataset in these formats. A Firefox extension called RDF Triple Collector (RTC) was developed, it can extract triple from web pages (annotated using RDFa), upload triples to Corese server and query data using SPARQL endpoint. A prototype of LDP 1.1 (Linked Data Platform) is implemented using RTC as data collector.

Besides, with the purpose of improving Corese query performance and carrying out research work on distributed environment, we proposed and developed a heuristic-based query planning method within Corese. The approach includes 3 main steps: 1) generate extended SPARQL query triple pattern Graph (ESG), 2) estimate the cost of ESG using pre-defined heuristics and cost models and 3) search ESG to find a good query plan and rewrite the SPARQL query. The approach was evaluated using BSBM benchmark, the results suggest that the developed method optimized 60% of the query execution time averagely [77].

6.13. Sociocultural Ontologies

Participants: Papa Fary Diallo, Olivier Corby, Isabelle Mirbel.

6.13.1. Sociocultural Ontology : Upper-level and Domain Ontologies

We propose a process of sociocultural ontology development in order to promise and preserve the culture of a country through sharing the customs and history of different localities. This can be compared with the construction of a platform straddling ”corporate memory” and a ”social network”, but applied in the context of a country. This process is based on the Vygotskian Framework, a theory of Russian psychologist Lev Vygotsky. We worked on an upper-level ontology and mapped it on the Linked Open Data (LOD) cloud. We designed a sociocultural domain ontology for the Senegalese context and the platform design on top of Semantic Mediawiki (SMW). This allows Senegalese communities to share and co-construct their sociocultural knowledge. This work is published in [59].

6.13.2. Human Time Ontology

In the second step of the PhD thesis of P. F. Diallo, we focus on the consideration of the time in the modeled knowledge. The main objective of this work is to provide a vocabulary (ontology) to handle temporal information on semantic data. Thus, the first step was to create a meta-language which handles temporal knowledge representation in the socio-cultural field that can be used in a wider area. This meta-language allows 1) to model cyclic knowledge (non-convex interval), 2) knowledge about calendar, 3) convex intervals 4) modeling absolute and relative time, 5) modeling relations between intervals, 6) distinction between open and closed intervals, 7) concepts such as time stamps and 8) to set different time granularities. The second step is to propose an RDFS representation of this meta-language. Thus this representation, Human Time Ontology (HuTO), allows us to model complex time statement which are a date, an interval (convex and non-convex), relative and absolute time. HuTO allows also temporal data annotation which is the representation of temporal notions on knowledge (expressed as RDF triple) and allow to reason over it.

HuTO allows us to use a resource as a temporal marker for dating another resource. Our ontology allows the relative dating which is to determine the relative order of resources, without necessarily determining their absolute time. A major contribution of HuTO is the modeling of non-convex intervals but also requests that can treat all types of intervals. For temporal annotation data, HuTO provides an approach that can link two models: one for temporal information and another for knowledge of the modeled area. This approach facilitates the information retrieval when it is on temporal or non-temporal data. Thus HuTO can annotate resources, triples or named graphs.

ZENITH Project-Team

6. New Results

6.1. Highlights of the Year

- Patrick Valduriez received the 2014 Innovation Prize from Inria – Académie des sciences – Dassault Systems.
- Miguel Liroz-Gistau received the best presentation award from the Grid5000 Spring School 2014 in Lyon for his talk on “Using Grid5000 for MapReduce Experiments”.
- Triton, a new common lab. (i-lab) has been created between Zenith and Beepeers (beepeers.com) to work on a platform for developing social networks in mobile/Web environments.
- 127 research groups worldwide registered to the LifeCLEF 2014 evaluation campaign chaired by Alexis Joly.

6.2. Big Data Integration

6.2.1. Probabilistic Data Integration

Participants: Reza Akbarinia, Naser Ayat, Patrick Valduriez.

Data uncertainty in scientific applications can be due to many different reasons: incomplete knowledge of the underlying system, inexact model parameters, inaccurate representation of initial boundary conditions, inaccuracy in equipments, error in data entry, etc.

An important problem that arises in big data integration is that of Entity Resolution (ER). ER is the process of identifying tuples that represent the same real-world entity. The problem of *entity resolution over probabilistic data* (which we call ERPD) arises in many distributed application domains that have to deal with probabilistic data, ranging from sensor databases to scientific data management. The ERPD problem can be formally defined as follows. Let e be an uncertain entity represented by multiple possible alternatives, i.e., tuples, each with a membership probability. Let D be an uncertain database composed of a set of tuples each associated with a membership probability. Then, given e, D, and a similarity function F, the problem is to find the entity-tuple pair (t, t_i) (where $t \in e, t_i \in D$) such that (t, t_i) has the highest cumulative probability to be the most similar in all possible worlds. This entity-tuple pair is called the *most probable match pair* of e and D, denoted by MPMP(e, D).

Many real-life applications produce uncertain data distributed among a number of databases. Dealing with the ERPD problem for distributed data is quite important for such applications. A straightforward approach for answering distributed ERPD queries is to ask all distributed nodes to send their databases to a central node that deals with the problem of ER by using one of the existing centralized solutions. However, this approach is very expensive and does not scale well neither in the size of databases, nor in the number of nodes.

In [24], we propose an efficient solution for the ERPD problem. Our contributions are summarized as follows. We adapted the possible worlds semantics of probabilistic data to define the problem of ERPD based on both similarity and probability of tuples. We proposed a PTIME algorithm for the ERPD problem. This algorithm is applicable to a large class of the similarity functions, where the similarity score of two tuples depends only on their attributes i.e., context-free functions. For the rest of similarity functions (i.e., context-sensitive), we proposed a Monte Carlo approximation algorithm. We also proposed a parallel version of our Monte Carlo algorithm using the MapReduce framework. We conducted an extensive experimental study to evaluate our approach for ERPD over both real and synthetic datasets. The results show the effectiveness of our algorithms.

Another topic of interest is the integration of large astronomy data catalogs. The main challenge in such integration, besides the huge amount of catalog data to be merged, is the weak identification of sky objects, which leads to ambiguities in object matching amongst catalogs. In cite [30], we present the NACluster algorithm. NACluster considers a Euclidian metric space and distance function to drive disambiguation amongst objects in various catalogs and extends the traditional k-means algorithm to deal with the dynamic creation of new clusters, representing real sky objects. NACluster shows F-measure results steadily superior to the Q3C join operator matching results, which is its closest competitor.

6.2.2. CloudMdsQL, a query language for heterogeneous data stores

Participants: Carlyna Bondiombouy, Boyan Kolev, Oleksandra Levchenko, Patrick Valduriez.

The blooming of different cloud data management infrastructures, specialized for different kinds of data and tasks, has led to a wide diversification of DBMS interfaces and the loss of a common programming paradigm. The CoherentPaaS European project addresses this problem, by providing a common programming language and holistic coherence across different cloud data stores.

In this context, we have started the design of a Cloud Multi-datastore Query Language (CloudMdsQL), and its query engine. CloudMdsQL is a functional SQLlike language, capable of querying multiple heterogeneous data stores (relational and NoSQL) within a single query that may contain embedded invocations to each data store's native query interface. Thus, CloudMdsQL unifies a quite diverse set of data management technologies while preserving the expressivity of their local query languages. Our experimental validation, with three data stores (graph, document and relational) and representative queries, shows that CloudMdsQL satisfies the five important requirements for a cloud multidatabase query language.

6.2.3. Semantic Data Integration using Bio-Ontologies

Participants: Emmanuel Castanier, Patrick Valduriez.

Biologist have adopted ontologies for several reasons: (1) to provide canonical representation of scientific knowledge; (2) to annotate experimental data to enable interpretation, comparison, and discovery across databases; (3) to facilitate knowledge-based applications for decision support, natural language processing and data integration. The challenge is to automatically process complex databases and generate mappings using relevant ontologies in a way that scales up for many resources and ontologies, while being easy to use for the biomedical community, customizable to fit specific needs and smart, in order to leverage the knowledge contained in ontologies.

The National Center for Biomedical Ontology (NCBO) has developped a popular ontology-based annotation workflow. To address the above challenge, we have integrated the NCBO annotator with our WebSmatch tool and the Biosemantic tool from IRD to perform semantic annotation using bio-ontologies [47]. The resulting tool provides very useful capabilities. First, it can convert SQL database schemas to RDF/RDFS with Biosemantic. Second, it can annotate with the NCBO annotator and WebSmatch using the NCBO resources index. Third, the NCBO annotator relies on WebSmatch to create mappings between elements of schemas and ontological concepts, and uses ontologies properties (i.e. subsomption, transitivity) to enhance matching techniques.

Unlike the bio-medical domain which has accepted ontologies as a means to manage (integrate) knowledge, the agronomic sciences is yet to exploit its full potential. To this end, we are currently developing an RDF knowledge base, Agronomic Linked Data (AgroLD) [50]. The knowledge base is designed to integrate data from various publically available plant centric data sources. The aim of AgroLD project is to collaborate with domain experts in bridging the gap between technology and its potential users to enhance biological research.

6.3. Distributed Indexing and Searching

6.3.1. Query Reformulation in P2P Data Management Systems

Participant: Esther Pacitti.

We consider peer-to-peer data management systems (PDMS), where each peer maintains mappings between its schema and some acquaintances, along with social links with peer friends. In this context, we deal with reformulating conjunctive queries from a peer's schema into other peer's schemas. Precisely, queries against a peer node are rewritten into queries against other nodes using schema mappings thus obtaining query rewritings. Unfortunately, not all the obtained rewritings are relevant to a given query, as the information gain may be negligible or the peer is not worth exploring. On the other hand, the existence of social links with peer friends might be useful to get relevant rewritings.

In [19], we propose a new notion of "relevance" of a query with respect to a mapping that encompasses both a local relevance (the relevance of the query wrt. the mapping) and a global relevance (the relevance of the query wrt. the entire network). Based on this notion, we design a new query reformulation approach for social PDMS which achieves great accuracy and flexibility. We combine several techniques: (i) social links are expressed as FOAF (Friend of a Friend) links to characterize peer's friendship; (ii) concise mapping summaries are used to obtain mapping descriptions; (iii) local semantic views are special views that contain information about mappings captured from the network by using gossiping techniques. Our experimental evaluation, based on a prototype on top of PeerSim and a simulated network demonstrate that our solution yields greater recall, compared to traditional query translation approaches proposed in the literature.

6.3.2. Diversified and Distributed Recommendation for Scientific Data

Participants: Esther Pacitti, Maximilien Servajean.

Recommendation is becoming a popular mechanism to help users find relevant information in large-scale data (scientific data, web). Different diversification techniques have been proposed to avoid redundancy in the process of recommendation. Intuitively, the goal of recommendation diversification is to identify a list of items that are dissimilar, but nonetheless relevant to the user's interests.

The main goal of this work [39], [17] is to define a new diversified search and recommendation solution suited for scientific data (i.e., plant phenotyping, botanical data). We first propose an original profile diversification scoring function that enables to address the problem of returning redundant items, and enhances the quality of diversification compared to the state-of-the-art solutions. We believe our work is the first to investigate profile diversity to address the problem of returning highly popular but too-focused items.Through experimental evaluation using two benchmarks we showed that our scoring function presents the best compromise between diversity and relevancy. Next, to implement our new scoring function, we propose a Top-k threshold-based algorithm that exploits a candidate list to achieve diversification. However this algorithm is greedy and does not scale up well.To overcome this limitation, we propose several techniques to improve performance. First, we simplify the scoring model to reduce its computational complexity. Second, we propose two techniques to reduce the number of items in the candidate list, and therefore the number of diversified scores to compute. Third, we propose different indexing scores (i.e., the score used to sort the items in the inverted lists) that take into account the diversification of items, and using them, we developed an adaptive indexing approach to reduce the number of accesses in the index dynamically based on the queries workload. We evaluated the performance of our techniques through experimentation. The results show that they enable to reduce the response time up to 12 times compared to a baseline greedy diversification algorithm.

We also address the problem of distributed and diversified recommendation (P2P and multi-site) that fits very well in different application scenarios. We propose a new scoring function (usefulness) to cluster relevant users over a distributed overlay. We analyzed the new clustering algorithm in details, and we studied its behavior with an experimental evaluation using different datasets. Compared with state-of-the-art solutions, we obtain major gains in recall (order of 3 times).

6.4. Scientific Workflows

6.4.1. Reuse of Scientific Workflows

Participant: Sarah Cohen-Boulakia.

With the increasing popularity of scientific workflows, public and private repositories are gaining importance as a means to share, find, and reuse such workflows. As the sizes of workflows repositories grow, methods to compare the scientific workflows stored in them become a necessity, for instance, to allow duplicate detection or similarity search. Scientific workflows are complex objects, and their comparison entails a number of distinct steps from comparing atomic elements to comparison of the workflows as a whole. Various studies have implemented methods for scientific workflow comparison and came up with often contradicting conclusions upon which algorithms work best. Comparing these results is cumbersome, as the original studies mixed different approaches for different steps and used different evaluation data and metrics.

We first contribute to the field [26] by (i) comparing in isolation different approaches taken at each step of scientific workflow comparison, reporting on an number of unexpected findings, (ii) investigating how these can best be combined into aggregated measures, and (iii) making available a gold standard of over 2000 similarity ratings contributed by 15 workflow experts on a corpus of 1500 workflows and re-implementations of all methods we evaluated.

Then, we present a novel and intuitive workflow similarity measure that is based on layer decomposition [40]. Layer decomposition accounts for the directed dataflow underlying scientific workflows, a property which has not been adequately considered in previous methods. We comparatively evaluate our algorithm using our gold standard and show that it a) delivers the best results for similarity search, b) has a much lower runtime than other, often highly complex competitors in structure-aware workflow comparison, and c) can be stacked easily with even faster, structure-agnostic approaches to further reduce runtime while retaining result quality.

6.4.2. Processing Scientific Workflows in Multi-site cloud

Participants: Ji Liu, Esther Pacitti, Patrick Valduriez.

As the scale of the data increases, scientific workflow management systems (SWfMSs) need to support workflow execution in High Performance Computing (HPC) environments. Because of various benefits, cloud emerges as an appropriate infrastructure for workflow execution. However, it is difficult to execute some scientific workflows in one cloud site because of geographical distribution of scientists, data and computing resources. Therefore, a scientific workflow often needs to be partitioned and executed in a multisite environment.

In [46], we define a multisite cloud architecture that is composed of traditional clouds, e.g., a pay-per-use cloud service such as Amazon EC2, private data-centers, e.g. a cloud of a scientific organization like Inria, COPPE or LNCC, and client desktop machines that have authorized access to the data-centers. We can model this architecture as a distributed system on the Internet, each site having its own computer cluster, data and programs. An important requirement is to provide distribution transparency for advanced services (i.e., workflow management, data analysis), to ease their scalability and elasticity. Current solutions for multisite clouds typically rely on application specific overlays that map the output of one task at a site to the input of another in a pipeline fashion. Instead, we define fully distributed services for data storage, intersite data movement and task scheduling.

Also, SWfMSs generally execute a scientific workflow in parallel within one site. In [38], we propose a non-intrusive approach to execute scientific workflows in a multisite cloud with three workflow partitioning techniques. We describe an experimental validation using an adaptation of Chiron SWfMS for Microsoft Azure multisite cloud. The experiment results reveal the efficiency of our partitioning techniques, and their superiority in different environments.

6.4.3. Data-centric Iteration in Dynamic Workflows

Participant: Patrick Valduriez.

Dynamic workflows are scientific workflows supporting computational science simulations, typically using dynamic processes based on runtime scientific data analyses. They require the ability of adapting the workflow, at runtime, based on user input and dynamic steering. Supporting data-centric iteration is an important step towards dynamic workflows because user interaction with workflows is iterative. However, current support for iteration in scientific workflows is static and does not allow for changing data at runtime.

In [20], we propose a solution based on algebraic operators and a dynamic execution model to enable workflow adaptation based on user input and dynamic steering. We introduce the concept of iteration lineage that makes provenance data management consistent with dynamic iterative workflow changes. Lineage enables scientists to interact with workflow data and configuration at runtime through an API that triggers steering. We evaluate our approach using a novel and real large-scale workflow for uncertainty quantification on a 640-core cluster. The results show impressive execution time savings from 2.5 to 24 days, compared to non-iterative workflow execution. We verify that the maximum overhead introduced by our iterative model is less than 5% of execution time. Also, our proposed steering algorithms are very efficient and run in less than 1 millisecond, in the worst-case scenario.

6.5. Scalable Query Processing

6.5.1. Big Data Partitioning

Participants: Reza Akbarinia, Miguel Liroz, Esther Pacitti, Patrick Valduriez.

The amount of data that is captured or generated by modern computing devices has augmented exponentially over the last years. For processing this *big data*, parallel computing has been a major solution in both industry and research. This is why, the MapReduce framework, which provides automatic distribution parallelization and fault-tolerance in a transparent way over lowcost machines, has become one of the standards in big data analysis.

For processing a big dataset over a cluster of nodes, one main step is data partitioning (or fragmentation) to divide the dataset to the nodes. In [23], we consider applications with very large databases, where data items are continuously appended. Thus, the development of efficient data partitioning is one of the main requirements to yield good performance. In particular, this problem is harder in the case of some scientific databases, such as astronomical catalogs. The complexity of the schema limits the applicability of traditional automatic approaches based on the basic partitioning techniques. The high dynamicity makes the usage of graph-based approaches impractical, as they require to consider the whole dataset in order to come up with a good partitioning scheme. In our work, we propose *DynPart* and *DynPartGroup*, two dynamic partitioning algorithms for continuously growing databases [23]. These algorithms efficiently adapt the data partitioning to the arrival of new data elements by taking into account the affinity of new data with queries and fragments. In contrast to existing static approaches, our approach offers constant execution time, no matter the size of the database, while obtaining very good partitioning efficiency. We validate our solution through experimentation over real-world data; the results show its effectiveness.

6.5.2. Scalable Query Processing with Big Data

Participants: Reza Akbarinia, Miguel Liroz, Patrick Valduriez.

We address the problem of data skew in MapReduce parallel processing framework. There are many cases where because of skew intermediate data, a high percentage of processing in the reduce side of MapReduce is done by a few nodes, or even one node, while the others remain idle.There have been some attempts to address this problem of data skew, but only for specific cases. In particular, there is no solution when all or most of the intermediate values correspond to a single key, or to a set of keys that are fewer than the number of reduce workers.

In this work, we propose FP-Hadoop, a system that makes the reduce side of MapReduce more parallel, and can efficiently deal with the problem of reduce side data skew. We extended the programming model of MapReduce to allow the collaboration of reduce workers on processing the values of an intermediate key, without affecting the correctness of the final results. In FP-Hadoop, the reduce function is replaced by two functions: intermediate reduce and final reduce. There are three phases, each phase corresponding to one of the functions: map, intermediate reduce and final reduce phases. In the intermediate reduce phase, the intermediate reduce function, which usually includes the main load of reducing in MapReduce jobs, is executed by reduce workers in a collaborative way, even if all values belong to only one intermediate key. This allows performing a big part of the reducing work by using the computing resources of all workers, even in the case of highly skewed data. We implemented a prototype of FP-Hadoop by modifying Hadoop's code, and conducted extensive experiments over synthetic and real datasets. The results show that FP-Hadoop makes MapReduce job processing much faster and more parallel, and can efficiently deal with skewed data. We achieve excellent performance gains compared to native Hadoop, e.g. more than 10 times in reduce time and 5 times in total execution time.

6.6. Data Stream Mining

6.6.1. Summarizing Uncertain Data Streams

Participants: Reza Akbarinia, Florent Masseglia.

In recent years, there has been a growing interest for probabilistic data management. In [41], we focus on probabilistic time series where a main characteristic is the high volumes of data, calling for efficient compression techniques. To date, most work on probabilistic data reduction has provided synopses that minimize the error of representation w.r.t. the original data. However, in most cases, the compressed data will be meaningless for usual queries involving aggregation operators such as SUM or AVG. We propose *PHA* (Probabilistic Histogram Aggregation), a compression technique whose objective is to minimize the error of such queries over compressed probabilistic data. We incorporate the aggregation operator given by the end-user directly in the compression technique, and obtain much lower error in the long term. We also adopt a global error aware strategy in order to manage large sets of probabilistic time series, where the available memory is carefully balanced between the series, according to their individual variability.

6.6.2. An Anti-Bouncing Data Stream Model

Participant: Florent Masseglia.

Usage mining is a significant research area with applications in various fields. However, Web usage data is usually considered streaming, due to its high volumes and rates. Because of these characteristics, we only have access, at any point in time, to a small fraction of the stream. When the data is observed through such a limited window, it is challenging to give a reliable description of the recent usage data. In [28] we show that data intralinkings, i.e., a usage record (event) may be associated with other records (events) in the same dataset, are common for Web usage streams. Therefore, in order to have a more authentic grasp of Web usage behaviors, the corresponding data stream models for Web usage streams should be able to process such intralinkings. We study the important consequences of the constraints and intralinkings, through the "bounce rate" problem and the clustering of usage streams. Then we propose the user-centric ABS (the Anti-Bouncing Stream) model which combines the advantages of previous models but avoids their drawbacks. First, ABS is the first data stream model that is able to seize the intralinkings between the Web usage records. It is also the first user-centric data stream model that can associate the usage records for the users in the Web usage streams. Second, owing to its simple but effective management principle, the data in ABS is available at any time for analysis. Under the same resource constraints as existing models in the literature, ABS can better model the recent data. Third, ABS can better measure the bounce rates for Web usage streams. We demonstrate its superiority through a theoretical study and experiments on two real-world data sets.

6.6.3. Autonomic Intrusion Detection: Adaptively Detecting Anomalies over Unlabeled Audit Data Streams

Participant: Florent Masseglia.

In [27], we propose a novel framework of autonomic intrusion detection that fulfills online and adaptive intrusion detection over unlabeled HTTP traffic streams in computer networks. The framework holds potential for self-managing: self-labeling, self-updating and self-adapting. Our framework employs the Affinity Propagation (AP) algorithm to learn a subject's behaviors through dynamical clustering of the streaming data. It automatically labels the data and adapts to normal behavior changes while identifying anomalies. Two large real HTTP traffic streams collected in our institute as well as a set of benchmark KDD'99 data are used to validate the framework and the method. The test results show that the autonomic model achieves better results in terms of effectiveness and efficiency compared to adaptive Sequential Karhunen-Loeve method and static AP as well as three other static anomaly detection methods, namely k-NN, PCA and SVM.

6.7. Scalable Data Analysis

6.7.1. Retrieval of Large-scale Visual Entities

Participants: Valentin Leveau, Alexis Joly, Patrick Valduriez.

In [37], we consider the problem of recognizing legal entities in visual contents in a similar way to named-entity recognizers for text documents. Whereas previous works were restricted to the recognition of a few tens of logotypes, we generalize the problem to the recognition of thousands of legal persons, each being modeled by a rich corporate identity automatically built from web images. We therefore introduce a new geometrically-consistent instance-based classification method that has several benefits over state-of-the-art instance classification methods including an efficient training phase reduced to a simple indexing process with a linear time and space complexity, but also the easy management of multi-labeled images, the fine grained localisation of the recognized patterns or the possibility of dynamically inserting additional training images in an incremental way. Experiments show that our method achieves better results than state-of-the-art techniques while being much more scalable, notably on an automatic web crawl of 5,824 legal entities which demonstrates the scalability of the approach.

6.7.2. Content-based Life Species Identification in Large Multimedia Collections

Participants: Alexis Joly, Julien Champ, Jean-Christophe Lombardo.

Building accurate knowledge of the identity, the geographic distribution and the evolution of living species is essential for a sustainable development of humanity as well as for biodiversity conservation. In this context, using crowdsourced data collection and multimedia identification tools is considered as one of the most promising solution. With the recent advances in digital devices/equipment, network bandwidth and information storage capacities, the production of multimedia data has indeed become an easy task. The emergence of citizen sciences and social networking tools has actually fostered the creation of large and structured communities of nature observers (e.g. e-bird, xeno-canto, Tela Botanica, etc.) who started to produce outstanding collections of multimedia records. Unfortunately, the performance of the state-of-the-art multimedia analysis techniques on such data is still not well understood and is far from reaching the real world's requirements in terms of identification tools. We therefore created LifeCLEF [36], [35], [31], [42], a new lab of the CLEF international forum [0] that evaluates these challenges in the continuity of the image-based plant identification task that we organized since 2011 within the ImageCLEF [0] lab. LifeCLEF is organized around 3 complementary tasks (PlantCLEF, BirdCLEF, FishCLEF), each being based on large and real-world data, as well as realistic scenarios established in collaboration with biologists and environmental stakeholders. 127 research groups worldwide did registered to the 2014 pilot campaign and downloaded the data. 22 of them crossed the finish line by submitting runs and papers to the workshop.

Besides the organization of the campaign, we also participated to two tasks in order to evaluate the content-based retrieval technologies developed within ZENITH. We notably implemented a new method [34] for the bird task based on the dense indexing of MFCC features and the offline pruning of the non-discriminant ones. To make such strategy scalable to the 30M of MFCC features extracted from the tens of thousands audio recordings of the training set, we used high dimensional hashing techniques coupled with an efficient approximate nearest neighbors search algorithm with controlled quality. Further improvements were obtained by (i) using a sliding classier with max pooling, (ii) weighting the query features according to their semantic coherence, and (iii) making use of the metadata to filter incoherent species. Results did show the effectiveness of the proposed technique which ranked 3rd among the 10 participating groups (some of them with years of experience in bioacoustic).

We finally investigated new interactive identification methods in [29], by extending classical faceted search mechanisms to the use of so called visual facets. The principle is to automatically build comprehensive visual illustrations of the expert data available in classical structured botanical dataset by building a visual matching graph of the related pictures and choosing the most connected ones. Additional facets can then be built automatically by clustering the graph and solving incompleteness issues.

[0]http://www.clef-initiative.eu/

[0]http://www.imageclef.org/

6.7.3. A look inside the Pl@ntNet experience

Participants: Alexis Joly, Julien Champ, Jean-Christophe Lombardo.

Pl@ntNet is an innovative participatory sensing platform relying on image-based plants identication as a mean to enlist non-expert contributors and facilitate the production of botanical observation data [22]. 18 months after the public launch of the iOS public application (and 6 months after the release of the Android version [32]), we carried out a self-critical evaluation of the experience with regard to the requirements of a sustainable and effective ecological surveillance tool (to appear in Multimedia Systems journal). Thanks to usage data analytics, we first demonstrated the attractiveness of the developed multimedia system (with more than 300K end-users and several thousands of users daily) as well as the nice self-improving capacities of the whole collaborative workflow (1.5 millions of observations were collected). We also pointed out the current limitations of the approach towards producing timely and accurate distribution maps of plants at a very large scale. We discussed in particular two main issues:

1. Data validation bottleneck: within the current workflow, only a few percentage of the observations are validated to avoid submerging the volunteer experts who actively do this job thanks to the collaborative web tools. There is consequently a need of smarter task assignment and recommendation mechanisms that would better balance the collaborative workload across all users and improves the serendipity.
2. Bias of the produced data: The temporal and geographical distribution of the observations is highly correlated with human activity. High densities of observations are more determined by population density and humans behavior than by plants density. This issue inevitably arises in any participatory sensing system but when the objective is to monitor noise nuisance or air quality, the concentration of the observations in the cities is less critical. There is therefore a need to build new data analytics methods compensating the bias through long-term statistics and the use of contextual information.