RESEARCH CENTER
Paris - Rocquencourt

FIELD

Activity Report 2012

Section New Results

Edition: 2013-04-24

ABSTRACTION Project-Team

6. New Results

6.1. Analysis of Biological Pathways

We have improved our framework to design and analyze biological networks. This framework focused on protein-protein interaction networks described as graph rewriting systems. Such networks can be used to model some signaling pathways that control the cell cycle. The task is made difficult due to the combinatorial blow up in the number of reachable species (*i.e.*, non-isomorphic connected components of proteins).

6.1.1. Semantics

Participants: Jonathan Hayman, Tobias Heindel [CEA-List].

Domain-specific rule-based languages can be understood intuitively as transforming graph-like structures, but due to their expressivity these are difficult to model in 'traditional' graph rewriting frameworks.

In [21], we introduce pattern graphs and closed morphisms as a more abstract graph-like model and show how Kappa can be encoded in them by connecting its single-pushout semantics to that for Kappa. This level of abstraction elucidates the earlier single-pushout result for Kappa, teasing apart the proof and guiding the way to richer languages, for example the introduction of compartments within cells.

6.1.2. Semantics and causality

Participants: Vincent Danos [University of Edinburgh], Jérôme Feret, Walter Fontana [Harvard Medical School], Russ Harmer [Paris VII], Jonathan Hayman, Jean Krivine [Paris VII], Chris Thompson-Walsh [University of Cambridge], Glynn Winskel [University of Cambridge].

In [20], we introduce a novel way of constructing concise causal histories (pathways) to represent how specified structures are formed during simulation of systems represented by rulebased models. This is founded on a new, clean, graph-based semantics introduced in the first part of this paper for Kappa, a rule-based modelling language that has emerged as a natural description of protein-protein interactions in molecular biology. The semantics is capable of capturing the whole of Kappa, including subtle side-effects on deletion of structure, and its structured presentation provides the basis for the translation of techniques to other models. In particular, we give a notion of trajectory compression, which restricts a trace culminating in the production of a given structure to the actions necessary for the structure to occur. This is central to the reconstruction of biochemical pathways due to the failure of traditional techniques to provide adequately concise causal histories, and we expect it to be applicable in a range of other modelling situations.

6.1.3. Case study: Combinatorial drift in yeast model

Participants: Vincent Danos [University of Edinburgh], Eric Deeds [University of Kansas], Jérôme Feret, Walter Fontana [Harvard Medical School], Russ Harmer [Paris VII], Jean Krivine [Paris VII].

The assembly of molecular machines and transient signaling complexes does not typically occur under circumstances in which the appropriate proteins are isolated from all others present in the cell. Rather, assembly must proceed in the context of large-scale protein-protein interaction (PPI) networks that are characterized both by conflict and combinatorial complexity. Conflict refers to the fact that protein interfaces can often bind many different partners in a mutually exclusive way, while combinatorial complexity refers to the explosion in the number of distinct complexes that can be formed by a network of binding possibilities.

In [9], we use computational models so as to explore the consequences of these characteristics for the global dynamics of a PPI network based on highly curated yeast two-hybrid data. The limited molecular context represented in this data-type translates formally into an assumption of independent binding sites for each protein. The challenge of avoiding the explicit enumeration of the astronomically many possibilities for complex formation is met by a rule-based approach to kinetic modeling. Despite imposing global biophysical constraints, we find that initially identical simulations rapidly diverge in the space of molecular possibilities, eventually sampling disjoint sets of large complexes. We refer to this phenomenon as “compositional drift". Since interaction data in PPI networks lack detailed information about geometric and biological constraints, our study does not represent a quantitative description of cellular dynamics. Rather, our work brings to light a fundamental problem (the control of compositional drift) that must be solved by mechanisms of assembly in the context of large networks. In cases where drift is not (or cannot be) completely controlled by the cell, this phenomenon could constitute a novel source of phenotypic heterogeneity in cell populations.

6.1.4. Automatic Reduction of Stochastic Semantics

Participants: Ferdinanda Camporesi, Jérôme Feret, Norman Ferns, Thomas Henzinger [Institute of Science and Technology, Austria], Heinz Koeppl [ETH Zürich], Tatjana Petrov [ETH Zürich].

Biology, Protein-protein interaction networks, Stochastic semantics, Verification.

We have proposed an abstract interpretation-based framework for reducing the state-space of stochastic semantics for protein-protein interaction networks. Our framework ensures that the trace distribution of the reduced system is the exact projection of the trace distribution of the concrete system. Moreover, when the abstraction is complete, if each state with the same abstraction is equiprobable at initial state, each state with the same abstraction is equiprobable at any time t.

In [10], we have formalized the model reduction framework for the stochastic semantics and we have established the relationships with the notions of lumpability, and bisimulation.

In [13], we have showed that the reduced models can be expressed in Kappa, and we have provided a procedure to do it.

6.2. Leakage Analysis

Participants: Matteo Zanioli [Correspondent], Pietro Ferrara [ETH, Zurich], Agostino Cortesi [Università Ca’ Foscari].

Abstract interpretation, Information leakage analysis, Object-oriented software, Static analysis.

In [28], we present SAILS, a new tool that combines SAMPLE, a generic static analyzer, and a sophisticated domain for leakage analysis. This tool does not require to modify the original language, since it works with mainstream languages like JAVA™, and it does not require any manual annotation. SAILS can combine the information leakage analysis with different heap abstractions, inferring information leakage over programs with complex data structures. SAILS has been applied to the analysis of the SecuriBench-micro suite. The experimental results underline the effectiveness of the analysis, since SAILS is in position to analyze several benchmarks in about 1 second without producing false alarms in more than 90% of the programs.

6.3. Termination

Participants: Patrick Cousot, Radhia Cousot.

Abstract interpretation, Computational induction, Induction, Proof, Static analysis, Semantic structural induction, Syntactic structural induction, Termination, Variant function, Verification.

In [17], we have introduced an abstract interpretation for termination.

Proof, verification and analysis methods for termination all rely on two induction principles: (1) a variant function or induction on data ensuring progress towards the end and (2) some form of induction on the program structure.

So far, no clear design principle did exist for termination as is the case for safety so that the existing approaches are scattered and largely not comparable with each other.

- For (1), we show that this design principle applies equally well to potential and definite termination. The trace-based termination collecting semantics is given a fixpoint definition. Its abstraction yields a fixpoint definition of the best variant function. By further abstraction of this best variant function, we derive the Floyd/Turing termination proof method as well as new static analysis methods to effectively compute approximations of this best variant function.
- For (2), we introduce a generalization of the syntactic notion of structural induction (as found in Hoare logic) into a semantic structural induction based on the new semantic concept of inductive trace cover covering execution traces by segments, a new basis for formulating program properties. Its abstractions allow for generalized recursive proof, verification and static analysis methods by induction on both program structure, control, and data. Examples of particular instances include Floyd's handling of loop cut-points as well as nested loops, Burstall's intermittent assertion total correctness proof method, and Podelski-Rybalchenko transition invariants.

6.4. Probabilistic Abstract Interpretation

Participants: Patrick Cousot, Michaël Monerau.

Abstract interpretation, Probabilistic systems, Static analysis.

Abstract interpretation has been widely used for verifying properties of computer systems. In [19], we present a way to extend this framework to the case of probabilistic systems.

The probabilistic abstraction framework that we propose allows us to systematically lift any classical analysis or verification method to the probabilistic setting by separating in the program semantics the probabilistic behavior from the (non-)deterministic behavior. This separation provides new insights for designing novel probabilistic static analyses and verification methods.

We define the concrete probabilistic semantics and propose different ways to abstract them. We provide examples illustrating the expressiveness and effectiveness of our approach.

6.5. Formal Verification by Abstract Interpretation

Participant: Patrick Cousot.

Abstract interpretation, Abstraction, Aerospace, Certification, Cyber-physical system, Formal Method, Mission-critical system, Runtime error, Safety-critical system, Scalability, Soundness, Static Analysis, Validation, Verification.

Abstract interpretation is a theory of abstraction and constructive approximation of the mathematical structures used in the formal description of programming languages and the inference or verification of undecidable program properties. Developed in the late seventies with Radhia Cousot, it has since then been considerably applied to many aspects of programming, from syntax, to semantics, and proof methods where abstractions are sound and complete but incomputable to fully automatic, sound but incomplete approximate abstractions to solve undecidable problems such as static analysis of infinite state software systems, contract inference, type inference, termination inference, model-checking, abstraction refinement, program transformation (including watermarking), combination of decision procedures, security, malware detection, etc.

This last decade, abstract interpretation has been very successful in program verification for mission- and safety-critical systems [12]. An example is ASTRÉE which is a static analyzer to verify the absence of runtime errors in structured, very large C programs with complex memory usages, and involving complex boolean as well as floating-point computations (which are handled precisely and safely by taking all possible rounding errors into account), but without recursion or dynamic memory allocation. Astrée targets embedded applications as found in earth transportation, nuclear energy, medical instrumentation, aeronautics and space flight, in particular synchronous control/command such as electric flight control or more recently asynchronous systems as found in the automotive industry. Astrée is industrialized by AbsInt Angewandte Informatik GmbH.

6.6. Static Analysis of Parallel Software

Participant: Antoine Miné.

Abstract interpretation, Embedded software, Parallel software, Rely/guarantee analysis, Run-time errors, Static analysis.

We present in [11] the theoretical foundation and the latest experimental evaluation of ASTRÉEA (5.3), a static analyzer prototype based on abstract interpretation to check for run-time errors in multi-threaded embedded critical C programs. Our method is based on a slightly modified non-parallel analysis that, when analyzing a thread, applies and enriches an abstract set of thread interferences. An iterator then re-analyzes each thread in turn until interferences stabilize. We prove the soundness of our method with respect to the sequential consistency semantics, but also with respect to a reasonable weakly consistent memory semantics. We also show how to take into account mutual exclusion and thread priorities through a partitioning over an abstraction of the scheduler state. This work is an extension of [54], complete with a full formalization and soundness proofs.

In [24], we express rely/guarantee methods in constructive form as an abstract interpretation of the interleaving trace semantics. We also restate the analysis presented in [11] as a further abstraction of rely/guarantee. This theoretical work brings a new understanding of the various causes of incompleteness and imprecision in our previous analysis, including the non-relational, input-insensitive, flow-insensitive, and history-insensitive treatment of interferences, and it opens the way to designing more precise analyses.

6.7. Static Analysis of Bit-Level Machine Integer and Floating-Point Operations

Participant: Antoine Miné.

Abstract interpretation, Embedded software, Numerical abstract domains, Run-time errors, Static analysis.

We present in [22] a few lightweight numeric abstract domains to analyze C programs that exploit the binary representation of numbers in computers, for instance to perform "compute-through-overflow" on machine integers, or to directly manipulate the exponent and mantissa of floating-point numbers. On integers, we propose an extension of intervals with a modular component, as well as a bitfield domain. On floating-point numbers, we propose a predicate domain to match, infer, and propagate selected expression patterns. These domains are simple, efficient, and extensible. We have included them into the ASTRÉE (5.2) and ASTRÉEA (5.3) static analyzers to supplement existing domains. Experimental results show that they can improve the analysis precision at a reasonable cost.

6.8. Inferring Sufficient Conditions with Backward Polyhedral Under-Approximations

Participant: Antoine Miné.

Abstract interpretation, Backward analysis, Numerical abstract domains, Static analysis, Sufficient condition inference, Under-approximations.

In [23], we discuss the automatic inference of sufficient pre-conditions by abstract interpretation and sketch the construction of an under-approximating backward analysis. We focus on numeric domains and propose transfer functions, including a lower widening, for polyhedra, without resorting to disjunctive completion nor complementation, while soundly handling non-determinism. A limited proof-of-concept prototype was designed to validate our approach. Planned applications include the derivation of sufficient conditions for a program to never step outside an envelope of safe states, or dually to force it to eventually fail.

6.9. A Constraint Solver Based on Abstract Domains

Participants: Marie Pelleau [University of Nantes, LINA], Antoine Miné, Charlotte Truchet [University of Nantes, LINA], Frédéric Benhamou [University of Nantes, LINA].

Abstract interpretation, Backward analysis, Numerical abstract domains, Static analysis, Sufficient condition inference, Under-approximations.

In [25], we apply techniques from abstract interpretation to constraint programming (which aims at solving hard combinatorial problems with a generic framework based on first-order logics). We highlight some links and differences between these fields: both compute fixpoints by iterations but employ different extrapolation and refinement strategies; moreover, consistencies in Constraint Programming can be mapped to non-relational abstract domains. We then use these correspondences to build an abstract constraint solver that leverages abstract interpretation techniques (such as relational domains) to go beyond classic solvers. We present encouraging experimental results obtained with our prototype implementation.

6.10. Automatic Inference of Necessary Preconditions

Participants: Patrick Cousot, Radhia Cousot, Manuel Fahndrich [Microsoft Research, Redmond, USA], Francesco Logozzo [Microsoft Research, Redmond, USA].

Abstract interpretation, Backward analysis, Static analysis, Necessary condition inference,

In [18], we consider the problem of automatic precondition inference for: (i) program verification; (ii) helping the annotation process of legacy code; and (iii) helping generating code contracts during code refactoring. We argue that the common notion of sufficient precondition inference (i.e., under which precondition is the program correct?) imposes too large a burden on call-sites, and hence is unfit for automatic program analysis. Therefore, we define the problem of necessary precondition inference (i.e., under which precondition, if violated, will the program always be incorrect?). We designed and implemented several new abstract interpretation-based analyses to infer necessary preconditions. The analyses infer atomic preconditions (including disjunctions), as well as universally and existentially quantified preconditions.

We experimentally validated the analyses on large scale industrial code.

For unannotated code, the inference algorithms find necessary preconditions for almost 64% of methods which contained warnings. In 27% of these cases the inferred preconditions were also sufficient, meaning all warnings within the method body disappeared. For annotated code, the inference algorithms find necessary preconditions for over 68% of methods with warnings. In almost 50% of these cases the preconditions were also sufficient. Overall, the precision improvement obtained by precondition inference (counted as the additional number of methods with no warnings) ranged between 9% and 21%.

6.11. Inference of Necessary Field Conditions with Abstract Interpretation

Participants: Mehdi Bouaziz, Manuel Fahndrich [Microsoft Research, Redmond, USA], Francesco Logozzo [Microsoft Research, Redmond, USA].

In [15], we present a new static analysis to infer necessary field conditions for object-oriented programs. A necessary field condition is a property that should hold on the fields of a given object, for otherwise there exists a calling context leading to a failure due to bad object state. Our analysis also infers the provenance of the necessary condition, so that if a necessary field condition is violated then an explanation containing the sequence of method calls leading to a failing assertion can be produced.

When the analysis is restricted to readonly fields, i.e., fields that can only be set in the initialization phase of an object, it infers object invariants. We provide empirical evidence on the usefulness of necessary field conditions by integrating the analysis into cccheck, our static analyzer for .NET.

Robust inference of readonly object field invariants was the #1 request from cccheck users.

6.12. TreeKs: A Functor to Make Numerical Abstract Domains Scalable

Participant: Mehdi Bouaziz.

Relational numerical abstract domains do not scale up. To ensure a linear cost of abstract domains, abstract interpretation-based tools analyzing large programs generally split the set of variables into independent smaller sets, sometimes sharing some non-relational information. In [14], we present a way to gain precision by keeping fully expressive relations between the subsets of variables, whilst retaining a linear complexity ensuring scalability.

6.13. An Abstract Domain to Infer Types over Zones in Spreadsheets

Participants: Cheng Tie, Xavier Rival.

abstract domains, spreadsheet script languages In [16], we proposed an abstract domain for the abstraction of spreadsheet contents.

Spreadsheet languages are very commonly used, by large user bases, yet they are error prone. However, many semantic issues and errors could be avoided by enforcing a stricter type discipline. As declaring and specifying type information would represent a prohibitive amount of work for users, we propose an abstract interpretation based static analysis for spreadsheet programs that infers type constraints over zones of spreadsheets, viewed as two-dimensional arrays. Our abstract domain consists in a cardinal power from a numerical abstraction describing zones in a spreadsheet to an abstraction of cell values, including type properties. We formalize this abstract domain and its operators (transfer functions, join, widening and reduction) as well as a static analysis for a simplified spreadsheet language. Last, we propose a representation for abstract values and present an implementation of our analysis.

6.14. Hierarchical Abstraction of Dynamic Structures

Participants: Pascal Sotin, Xavier Rival.

abstract domains, shape analysis, domain combination In [26], we designed a hierarchical shape abstract domain for the abstraction of complex data structures found in embedded softwares.

We propose a hierarchical shape abstract domain, so as to infer structural invariants of dynamic structures such as lists living inside static structures, such as arrays. This programming pattern is often used in safety critical embedded software that need to “allocate” dynamic structures inside static regions due to dynamic memory allocation being forbidden in this context. Our abstract domain precisely describes such hierarchies of structures. It combines several instances of simple shape abstract domains, dedicated to the representation of elementary shape properties, and also embeds a numerical abstract domain. This modular construction greatly simplifies the design and the implementation of the abstract domain. We provide an implementation, and show the effectiveness of our approach on a problem taken from a real code.

6.15. Reduced Product Combination of Abstract Domains for Shapes

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

abstract domains, shape analysis, reduced product In [27], we proposed a notion of reduced product for shape abstractions.

Real-world data structures are often enhanced with additional pointers capturing alternative paths through a basic inductive skeleton (e.g., back pointers, head pointers). From the static analysis point of view, we must obtain several interlocking shape invariants. At the same time, it is well understood in abstract interpretation design that supporting a separation of concerns is critically important to designing powerful static analyses. Such a separation of concerns is often obtained via a reduced product on a case-by-case basis. In this paper, we lift this idea to abstract domains for shape analyses, introducing a domain combination operator for memory abstractions. As an example, we present simultaneous separating shape graphs, a product construction that combines instances of separation logic-based shape domains. The key enabler for this construction is a static analysis on inductive data structure definitions to derive relations between the skeleton and the alternative paths. From the engineering standpoint, this construction allows each component to reason independently about different aspects of the data structure invariant and then separately exchange information via a reduction operator. From the usability standpoint, we enable describing a data structure invariant in terms of several inductive definitions that hold simultaneously.

AOSTE Project-Team

6. New Results

6.1. Logical time in Model-Driven Engineering embedded design

Participants: Charles André, Frédéric Mallet, Julien Deantoni, Marie-Agnès Peraldi Frati, Arda Goknil, Nicolas Chleq.

6.1.1. TimeSquare

We progressed our work on the foundations of logical time modeling as present in MARTE Time Model and our CCSL clock constraint specification language, while continuing the development of the TimeSquare tool environment which supports this in practice. A technical position paper was presented to the international TOOLS conference [22].

6.1.2. ECL (Event Constraint Language

Our contributions on CCSL and Time Model to the MARTE profile are part of the standard, but so far expressed in a syntax that is clearly distinct of the former UML notations. On the other hand, UML provides a textual language, named OCL, to express well-formedness constraints on diagram models and metamodels. While the original objectives were quite different, it seemed tempting to extend or adapt the general OCL philosophy, and to apply it then to timing and performance constraints as targeted by CCSL. The goal is to able the description of MoCs in an appropriate syntax, at metamodeling level. The result was a new syntax, called ECL for event contraint language, endowed with the well-established, sound timing interpretation as in CCSL. This work was reported in [40].

6.1.3. Logical time clocks to schedule data-flow models

Data-flow models can be used to capture data dependencies from applications, execution platforms and allocations. Most of the time such data dependencies impose only a partial order on the execution of application elements onto the execution platform and allow several allocation schemes. In [38], we have shown how to use logical time and CCSL constraints to capture explicitly the partial order imposed by the data-dependencies without imposing a total order. This work of representation expressivity then paved the way for analysis studies on time refinement, described in 6.3 .

6.1.4. Timing requirement modeling

One of the weak points of UML regarding a complete system design flow is its poor treatment of requirement capture (although this is partly corrected in the SysML profile). When requirements are made on timing aspects and logical time (as in our advocated approach), the relevant syntactic expressivity must be provided. We worked on the definition of a Domain-Specific Language (DSL for Timing Requirements engineering. The results were presented in [24], then applied to system specification in the context of the work described in section 6.6 .

6.2. Semantic translation of CCSL constraints into appropriate Büchi automata for trace recognition

Participants: Frédéric Mallet, Julien Deantoni, Robert de Simone, Ling Yin.

Our CCSL language expresses timing and scheduling constraints for a system, based on the notion of abstract logical clocks providing time events, and contraints linking them with relations of "asynchronous" nature (precedence, faster than) or of "synchronous" origin (subclocking, included in). Of course in a large system design both types coexist, and functional definitions also live next to declarative specifications to allow several timing solutions. Such a solution, called a schedule, must enforce that each logical clock either ticks endlessly, or terminates properly, in a way that globally respects the constraints. In previous works we have shown how a large variety of semantic scheduling constraints from the literature could soundly be represented in CCSL.

This year we focused on the semantic foundation of our CCSL language, by defining a structral operational semantic translation into a specific type of transition systems. Because we deal with infinite traces we had to consider acceptance mechanisms such as Büchi repeated states (as already used for translation of LTL temporal logic formulae in classical model-checking). Next we found out that, while state-labeled acceptance conditions were fine to obtain a direct and intuitive translation of individual constraints, building the composition of such models when dealing with multiple constraints was much easier in the case of *transition-labeled* Büchi automata (with repeated acceptance criteria now on transitions); the theory carries over to such case quite naturally, and has already been studied in the past. Finally, because traces must include infinite occurrences for *each* clock, we had to move to so-called *extended* Büchi automata, again a model already studied previously. We provided a complete semantic translation for all CCSL kernel constructs. Most importantly, we provided an efficient and simple fix-point algorithm to check the existance of a valid schedule, based on the type of automata just defined. This is (we believe) a genuine improvement on existing results, with potential applications outside our direct scope. These results are presented in a technical report, submitted for publication [46].

6.3. Timing refinement for multidimensional dataflow models using MARTE Time Model

Participants: Frédéric Mallet, Julien Deantoni, Jean-Vivien Millo.

Extensions of dataflow process networks have been proposed (as multidimensional SDF) to combine task parallelism (as in traditional process networks) with intensive data parallelism (as proposed in the Array-OL/Gaspard2 formalism developed in the DaRT EPI, for instance). The prospect of scheduling (seen as precise time cycle allocation) is here more complex, because of possible trade-offs between the granularity of treatments at task level *vs.* the size of data arrays that are handled uniformally in parallel inside each task. We considered how these phenomena could be represented (if not solved) inside the framework of MARTE Time Model and logical clocks, so as to handle such design issues in a well-defined MDE approach. Additionally, we used the MARTE platform description to specify how the previous models are refined through mapping allocation. The resulting modeling framework was presented in a journal article [19]. This work was conducted jointly with P. Boulet, from DaRT EPI, and C. Glitia, former DaRT PhD and Aoste postdoc student.

6.4. Process Network analysis

Participants: Robert de Simone, Jean-Vivien Millo.

6.4.1. K-periodic routing schemes for Network-on-Chip data traffic

This year we considered more specifically the issue of exploiting the predictable routing schemes of our KRG models, expressed as infinite binary words to indicate the sucessive branching directions at merge/select switch nodes, in order to encode data traffic patterns expanded at compile time, when mapping applications expressed under the form of dataflow process networks onto processor arrays in manycore architectures based on network-on-chip interconnects. To show the potential impact of such predicatble compile-time routing patterns, we stdudied as a typical example a fulll (all-to-all) broadcast algorithm on a mesh topology, connecting mode-less computation nodes as in the theory of cellular automata. This resulted in a precise recursive definition of routing patterns, which achieve an optimal data propagation (broadcast implemented as multicast), given the availability of actual links in the NoC topology. This result was presented at the Autamata'2012 conference [30], and an expanded version is available as technical report [44].

A wider view of the approach, and its potential benefits, are described in a technical report [43], submitted for publication.

6.4.2. Optimal data placement for process network scheduling

The topic of efficient scheduling of dataflow process network traffic to optimize both throughput and buffer queue sizing has given rise to a huge literature starting with seminal works in [49], [47], [56]. It has recently been given new impulse due to the advent of manycore architectures (see above). We conducted a number of theoretical works, to establish how such optimal computation scheduling can be best achieved in configurations where data are evenly distributed and streched in time across the (process) network. While this result is intuitively obvious, we formalized precisely what evenly distributed technically means, with the notion of balanced/mechanical words going a long way back in formal language theory, and we demonstrated that under such assumptions optimal schedules could be constructed *in a fully analytical way*, without any symbolic simulation steps or behavior expansion. The result was accepted for publication in a journal article [20].

6.5. Transformation from MARTE Time Model and CCSL to formal analysis models

Participants: Frédéric Mallet, Ling Yin.

This work was conducted in the context of an on-going collaboration with the Software Engineering Institute (SEI) of East Normal China University (ECNU) at Shanghai, which led altogether in part to the DAESD Associated-team, followed by a LIAMA joint project proposal recently submitted (HADES), and the co-supervision by Frédéric Mallet (together with Professor Jing Liu from ECNU) of the PhD thesis of Yin Ling. Yin Ling spent a one-year visit in our team, funded on a chinese governemental grant.

We studied the efficient and sound formal translation of a subset of CCSL contraints into the PROMELA/SPIN formalism, to benefit from model-checking formal analysis features in this environment. The translation is not completely direct, as synchronous simultaneity is not a native notion of PROMELA, and has to be encoded as atomicity. The motivating principles and translation details are provided in [42]. A similar attempt could be considered in the future, this time with the synchronous model-checker SMV, which allows coumpound instantaneous atomic behaviors.

Another line of research was initiated at ECNU to consider *logical continuous time*, while most of our current work considers only discrete time (while MARTE Time Model considers both). Considerations on *hybrid state diagrams* , inviting the expressive power of formal hierarchical hybrid automata models into the MDE design space of UML MARTE, were investigated in [27].

6.6. Use of MARTE Time Model and Logical Time in automotive design and AUTOSAR/TADL

Participants: Marie-Agnès Peraldi Frati, Julien Deantoni, Arda Goknil.

Precise timing constraint modeling and analysis [26], [33] is a key point for the correct development of automotive electronics. EAST-ADL and AUTOSAR has been adopted as standards in automotive industry. The timing model (TADL :Time augmented Description Language) of these standards raises different issues, mainly concerning the precise modeling of the multi clock characteristics of distributed systems together with parameterized timing expressions. In the ITEA TIMMO-2-USE project [35] 8.3.2.1 , we conducted a work [34], [35], on extending TADL with an explicit notion of multiple time bases for modeling the various temporal referentials used in an automotive design (clocks from different ECUs, motor position, etc). Additionally, timing constraints are augmented with parameters, which can be free at the highest abstraction level and then progressively defined during the design process. As a result, a symbolic timing expression in TADL2 is possibly made of a suitable set of arithmetic operators mixing symbolic identifiers (not necessarily set variables) and referring to different time bases. One typical use of this feature is to capture unknown configuration parameters for time budgeting; another one is to relate constraints in different time-bases to each other. Inherent to this work is also the study of the allowable ranges for symbolic values that are dictated by a set of constraints.

6.7. Multiview modeling and power intent in Systems-on-chip

Participants: Carlos Gomez Cardenas, Ameni Khecharem, Jean-François Le Tallec, Frédéric Mallet, Julien Deantoni, Robert de Simone.

6.7.1. High-level power management modeling

One of the concern of the UML MARTE profile is to allow non-functional property modeling, so that the same system bare description can be annotated in a number of views. In our case, combined with our logical time framework, such properties can be made as time-depending, inside potentially distinct views. We examplified this approach by dealing to a large extent with the example of low-power design and energy modeling in the case of Systems-on-Chip (SoC) in the mobile phone domain. Pure power/thermal modeling can be realized, based on the system global architecture, then made operational with the use of logical time controllers triggering power management functionalities.
Thermal/power simulation models are usually relying on continuous time. Therefore we considered the issue of *logical continuous* time, in an early attempt at combining simulation of continuous time power/thermal models with intrinsically discrete functional aspects. A prototype was realized in Scicos, as part of Ameni Khecharem master internship.
This work was conducted in the context of Carlos Gomez PhD thesis, and in collaboration with several partners inside the ANR HeLP project. It should be continued in the forthcoming PhD thesis of Ameni Khecharem, just started in the context of the follow-up ANR HOPE project, which will consider specific issues of hierachical power modeling and compositional power management (as an example of incremental multiview aspects).

6.7.2. IP-XACT

In this context of high-level power modeling and multiview concerns, we considered the emerging Accelera standard IP-XACT, made to provide easy-to-plug interfaces and Architecture Description Language (ADL) to allow simple assembly of hardware IP components into well-behaved SoCs. More specifically we provided means to annotate such interface with extra informations, directly borrowed from UML MARTE NFP properties, to handle power and thermal aspects. A number of model transformations back and forth between MARTE and (extended) IP-XACT were realized, and extraction of IP-XACT compliant interfaces from proprietary SystemC code describing the elementary IP component tehmselves has been defined and implemented as well.
This work was initiated as part of a project with STMicroelectronics, inside the nano2012 programme (ended 2011), and continued as part of the ANR HeLP collaboration. It resulted in the PhD thesis of Jean-François Le Tallec (who remained in the team for a couple of months later to complete the prototype implementation) [16].

6.8. Correct and efficient implementation of polychronous formalisms

Participants: Thomas Carle, Manel Djemal, Dumitru Potop Butucaru, Robert de Simone, Yves Sorel.

We extended our work on extending the AAA methodology for polychronous processes, by providing a better integration of clock analysis in the various phases of the implementation process (allocation, scheduling, pipelining, etc.). We also considered a wider range of implementation targets (time-triggered, MPSoC) and non-functional constraints (partitioning).

6.8.1. Time-Triggered Platform targets

Our first result this year concerns the automatic scheduling and code generation for time-triggered platforms. We extended our previous results in two significant ways. First, we designed a novel approach for specification of real-time features of time-triggered systems, with deadlines longer than periods; this allows a faithful representation of complex end-to-end flow requirements. Second, we provided new algorithms for off-line pipelined scheduling of these specifications onto partitioned time-triggered architectures *à la* ARINC 653; allocation of time slots/windows to partitions can be either complete or partially provided, or synthesized by our tool. Automatic allocation and scheduling onto multi-processor (distributed) systems with a global time base becomes feasible, taking into account communication costs. For single processors, we allow the generation of fully compliant ARINC653/APEX implementation code.

This work was mainly carried out inside the FUI Parsec 8.2.2.2 (which funds the PhD thesis of T. Carle) and P 8.2.2.1 projects, as well as a collaboration with ASTRIUM Space Transportation. First results are presented in a technical report, submitted for publication [39].

6.8.2. Multi-Processor System-on-Chip (MP-SoC) targets

Our second contribution concerns the automatic allocation and real-time scheduling over MPSoC (multi-processor on chip) architectures with NoC (network-on-chip) interconnect. One must take into account the specific 2D mesh network-on-chip topology, and synthesize the NoC routing patterns. This work provides operational execution support for the contributions described in 6.9 .

6.8.3. The LoPhT tool

Our recent work on extending the AAA methodology with better handling of execution conditions, with pipelining and pipelined scheduling, and with specific real-time scheduling and code generation techniques for time-triggered/partitioned and MPSoC platforms resulted in the development of a new scheduling and code generation toolbox, called LoPhT (for Logical to Physical Time Compiler).

6.9. Programmable On-Chip Networks

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

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

Efficient compilation of applications onto MPSoCs remains largely an open problem, with the issue of best mapping of computation parts (threads, tasks,...) onto processing resources amply recognized, while the issue of best use of the interconnect NoC to route and transfer data still less commonly tackled. In the most general case, dynamic allocation of applications and channel virtualization can be guided by user-provided information under various forms, as in OpenMP, CUDA, OpenCL and so on. But then there is no clear guarantee of optimality, and first attempts by non-experts often show poor performances in the use of available computing power. Conversely there are consistent efforts, in the domains of embedded and HPC computing, aiming at automatic parallelization, compile-time mapping and scheduling optimization. They rely on the fact that applications are often known in advance, and deployed without disturbance from foreign applications, and without uncontrolled dynamic creation of tasks. Our contribution follows this “static application mapping” approach.

An optimal use of the NoC bandwidth should authorize data transfers to be realized according to (virtual) channels that are temporarily patterned to route data “just-in-time”. Previous works have identified the need for Quality of Service (QoS) in “some” data connections across the network (therefore borrowing notions from macroscopic networks, say internet and its protocols). But our experience with the AAA methodology strongly suggests that optimal NoC usage should result from a global optimization principle (embodied in a form of the AAA methodology), as opposed to a collection of local optimizations of individual connections. Indeed, various data flows with distinct sources and targets will nevertheless be highly concerted, both in time and space, like in a classical pipelined CPU, where the use of registers (replaced in our case with a complex NoC) is strongly synchronized with that of the functional units.

One main problem in applying such a global optimization approach is to provide the proper hardware infrastructures allowing the implementation of optimal computation and communication mappings and schedules. Our thesis is that optimal data transfer patterns should be encoded using simple programs configuring the router nodes (each router being then programmed to act its part in the global concerted computation and communication scheme).

We addressed this problem in the framework of our collaboration with the "Embedded Systems- on-Chips" department of the LIP6 laboratory, one of the main site of expertise for SoC/NoC design and Hardware/software codesign. This collaboration first materialized with the co-supervision of M. Djemal's PhD thesis. We concretely supported our proposed approach by extending the DSPIN 2D mesh network-on-chip (NoC) developed at UPMC- LIP6. In this NoC, we replace the fair arbitration modules of the NoC routers with static, micro-programmable modules that can enforce a given packet routing sequence, as specified by small programs. The design of such simple routing schemes can, for instance, be extracted from our results in section 6.4 .
We advocate the desired level of expressiveness/complexity for such simple configuration programs, and provide experimental data (cycle-accurate simulations) supporting our choices. We also wrote an architecture synthesis tool that allows simple architectural exploration of MPSoCs using the new DSPINPro NoC. First results in this direction have been presented in the DASIP 2012 conference, where our paper [23] has been short-listed for best paper award.

6.10. Uniprocessor Real-Time Scheduling

Participants: Laurent George, Mohamed Marouf, Daniel De Rauglaudre, Yves Sorel.

6.10.1. Combination of Non-Preemptive and Preemptive Tasks

We focused on fixed priority scheduling for a combination of non-preemptive strict periodic tasks in conjunction with preemptive sporadic tasks, that we extended to software fault tolerance [29]. We first investigated the transient phase for non-preemptive strict periodic tasks and we proved that its length is smaller than the transient phase for preemptive periodic tasks. Then, we determined the worst case scenario for preemptive sporadic tasks where the Worst Case Response Time (WCRT) can be obtained in the presence of strict periodic tasks. We proved that these release times belong only to the permanent phase of strict periodic tasks, and thus that the schedulability analysis for sporadic tasks can be restricted to the permanent phase. For preemptive sporadic tasks, we extended the classical necessary and sufficient schedulability condition based on the worst case response time computation to take into account non-preemptive strict periodic tasks. Finally, we considered software fault tolerance in the particular case where each primary strict periodic task has an alternate sporadic task which is released when the primary task fails. The schedulability analysis guarantees that even if all strict periodic tasks fail then all their respective alternate tasks will meet their deadlines.

6.10.2. Formal Proofs of Real-Time Scheduling Theorems

We completed two formal proofs of theorems in Coq on scheduling of fixed priority real-time preemptive tasks: one dealing with the sizes of busy periods (about 3500 lines of Coq), and another one dealing with response time (about 5200 lines of Coq). A monograph about these proofs, together with the formal check in Coq of scheduling conditions of strict periodicity, presented in the conference JFLA 2012 [37], have been started (currently about 70 pages).

6.11. Multiprocessor Real-Time Scheduling

Participants: Abderraouf Benyahia, Laurent George, Mohamed Marouf, Falou Ndoye, Simon Nivault, Yves Sorel, Cécile Stentzel, Meriem Zidouni.

6.11.1. Non-Preemptive Partitioned Fault Tolerant Scheduling

We addressed partitioned multiprocessor scheduling of non-preemptive strict periodic tasks which is extended thereafter to hardware fault tolerance [17].

In order to schedule a task set of non-preemptive strict periodic tasks on a multiprocessor platform, we partitioned this task set into subsets of tasks, each one is scheduled on a single processor using our proposed uniprocessor scheduling algorithm. The partition is carried out according to an enhanced "First Fit" algorithm that balances the load of the tasks on all the processors. However, inter-processors communications can lead to delay task execution. Thus, we determined the start time of each task taking into account the communication delay between this latter task and its predecessor tasks. Also, as inter-processor communications may generate a transient phase, we computed the length of the transient phase.

We proposed a fault tolerant real-time scheduling algorithm which allows hardware processors and/or buses faults, and conserves the strict periodicity of each task. We also proposed a graph transformation algorithm, applied on the task graph, which generates redundancies of tasks as well as dependencies. The transformation adds also selector tasks which choose data coming from the non failing processors and buses. That algorithm is based on exclusion relations to assign redundant tasks (resp. dependencies) to different processors (resp. busses). Then, we extended the previous partitioned multiprocessor scheduling algorithm to manage fault tolerance taking into account these exclusion relations.

This approach was successfully implemented on a CyCabs electric vehicle in a real-time fault tolerant tracking application where some processor or some bus could fail without any consequence on the proper execution of the application, i.e. same functional behaviour and real-time constraints satisfied.

6.11.2. Partitioned Scheduling with Exact Preemption Cost

Preemption allows a better scheduling success ratio but has a cost that must not be neglected in safety critical applications of domains such as avionic, automotive, etc. We focused on partitioned multiprocessor scheduling of independent preemptive periodic real-time tasks, while taking into account the exact preemption cost with the $\oplus$ operation formerly proposed by Meumeu and Sorel [10]. We improved the "greedy" heuristic proposed last year and compared it with the "Best-Fit" (BF) and "Worst-Fit" (WF) heuristics classically used in partitioned multiprocessor scheduling, but extended to take into account the exact preemption cost. We also compared our heuristic with an exact "Branch and Bound" algorithm with the same extension. The first comparison shows that the task allocation found by our heuristic gives a better response time than those found by WF and BF. This is due to the fact that the execution of the tasks is better parallelized. On the other hand, BF and WF heuristics execute a bit faster than our heuristic because they do not use all the available processors contrary to our heuristic which has the advantage to improve the load balancing of the tasks on all the processors.

Then, we addressed the scheduling of preemptive periodic real-time tasks with dependence constraints involving task precedences and data dependences. We considered harmonic tasks, i.e. periods of tasks are multiple or equal, to avoid loss of data. In order to satisfy data dependence constraints, we modified the release dates and deadlines of the dependent tasks according to the reception date of the data. In addition, data dependences between tasks mean to share data between dependent tasks which can cause deadlock and priority inversion problems. In order to solve these problems while taking into account the preemption cost, we proposed a new schedulability condition based on an extension of the $\oplus$ operation. We plan to propose a mutiprocessor scheduling heuristic based on that condition applied on tasks with modified release dates and deadlines.

6.11.3. Semi-partitioned Scheduling

Semi-partitioned multiprocessor scheduling stands between partitioned and global scheduling, the latter allowing migrations. We mainly addressed the semi-partitioned scheduling approach where the Worst Case Execution Time (WCET) of a job can be portioned, each portion being executed on a dedicated processor, according to a static pattern of migration. A job is migrated at its local deadline, computed from the deadline of the task it belongs to. We have studied this approach in the context of a fork/join task model with thread parallelism. A task is composed of a sequence of segments that can be parallelized in threads, if needed. The local deadlines depends on the number of parallel threads assigned to each segment.

6.11.4. Code Generation for Multicore

This work was carried out in the OPENPROD ITEA project 8.3.2.2 . xMod developed by IFPEN (IFP Energies Nouvelles), is an heterogeneous model integration environment that allows model importation from specific tools such as Simulink, AMSIM, etc. It also provides as a virtual instrumentation laboratory. In order to make xMod being able to run simulations with hardware-in-the-loop environment, we developed a new SynDEx executive kernel based on the kernel, dedicated to Windows/RTX, developed last year. That executive kernel is used with the macro-code generated by SynDEx to produce a real-time executable code that can drive the execution (real-time multi-core distribution and synchronized execution) of the models imported by xMod

and simulated in the virtual instrumentation laboratory. This prototype as well as the report describing the corresponding achieved works, are the final deliverable of the OPENPROD project.

Furthermore, a French and English SynDEx code generation reference manual has been written to help future SynDEx users and maintainers to generate real-time code for already supported architectures or new ones.

6.11.5. Gateway with Modeling Languages for Certified Code Generation

This work was carried out in the P FUI project 8.2.2.1 . We provide inside the project expertise mainly on schedulability analysis and automatic generation of distributed real-time code. In this context, we developed a gateway between UML/MARTE and SynDEx. From a model specified with UML (Activity Diagram to specify algorithms and Composite Structure Diagram to specify multicomponent architectures) and refined with the UML profile MARTE (Modeling and Analysis of Real-Time Embedded Systems), we use the gateway to generate automatically distributed real-time application specified in the SynDEx format. Currently, we intend to provide a gateway between the GeneAuto language and SynDEx. The GeneAuto language is a subset of the future pivot P language. We presently deal with the part of the GeneAuto language corresponding to Simulink for data-flow modeling and we plan to deal soon with the part corresponding to Stateflow for control-flow modeling (composition of automata).

6.11.6. SynDEx Updates

We continued the software developments for the future version 8 of SynDEx which will feature a new software architecture to allow better functionality evolutions and maintenance. On the other hand in the COTROS ADT ("Génération de code temps réel distribué optimisé et sûr"), we completed the tests on the new automatic code generator for the current version 7 of SynDEx. This new generator produces code for mono-periodic and multi-periodic applications with condition and repetitive control structures, for the different hardware architectures supported by SynDEx. We developed a checker for the generated code that was integrated in the new generator. This checker verifies the correct use of semaphores and consequently the absence of deadlocks in the real-time code. Deadlocks are the most difficult part when dealing with distributed architectures. We achieved also a maintenance report describing the structure and the main features of code generator, as well as the technical choices we did.

6.12. Variability of program execution times on multicore processors

Participants: Sid-Ahmed-Ali Touati, Matias Vara Larsen, Abdelhafid Mazouz.

The activity described here represents the finalization of previous efforts conducted by Sid Touati and members of his groups, initiated before he joined the AOSTE EPI, and which are progressively merged with our own objectives, for results to be reported hopefully next year).

With the massive introduction of multicore platforms on embedded systems, parallel applications gained in performance. However, we showed in previous studies that the performance gain comes with high instability: program execution times vary in important way. We investigated the reasons for this variations and tried to understand the factors that influence program performance variability, that we decompose intro multiple families: factors from the application itself (implemented algorithms, coding technique, synchronization barriers, etc.), factors from the execution environment (OS effects, thread scheduling, Input/Ouput operations) and factors from the underlying hardware (micro-architecture, memory hierarchy, speculative execution, hardware data prefectching, etc.). Now, we have better understanding to these factors thanks to the work of two students:

1. Mr. Abdelhafid Mazouz who defended his PhD under the direction of Sid Touati at the university of Versailles in 11th of December 2012. The title of his PhD is "An Empirical Study of Program Performance of OpenMP Applications on Multicore Platforms".
2. Mr. Matias Vara Larsen, intern under the supervision of Sid Touati from February to June 2012, inside the Aoste EPI in Sophia-Antipolis, co-funded under a grant from Inria international internship program). The topic of his internship was to study the influence of he Linux kernels (multiple versions) on the stability of parallel applications.

Last, we published a rigorous statistical protocol in [21] called the Speedup-Test. It is used to analyze valid speedups (performance gain) in presence of performance instability: The Speedup-Test protocol is implemented and distributed as an open source tool based on R software. Our statistical methodology defines a consistent improvement compared with the usual performance analysis method in high-performance computing.

CASCADE Project-Team (section vide)

CONTRAINTES Project-Team

6. New Results

6.1. Inferring Reaction Rule Models from Ordinary Differential Equations

Participants: François Fages, Steven Gay, Sylvain Soliman.

Many models in Systems Biology are described as Ordinary Differential Equations (ODEs), which allow for numerical integration, bifurcation analyses, parameter sensitivity analyses, etc. However, before fixing the kinetics and parameter values and going to simulations, various analyses can be performed based only on the structure of the model. This approach has rapidly developed in Systems Biology in the last decade, with for instance, the analyses of structural invariants in Petri net representation, model reductions by subgraph epimorphims, qualitative attractors in logical dynamics or temporal logic properties by analogy to circuit and program verification. These complementary analysis tools do not rely on kinetic information, but on the structure of the model with reactions.

In [8], [19], we present a symbolic computation algorithm for inferring a reaction model from an ODE system, based a general compatibility condition between the kinetic expression and the structure of a reaction, and report on its use for automatically curating the writing in SBML of the models in the respository biomodels.net. SBML is now a standard for sharing and publishing reaction models. However, since SBML does not enforce any coherence between the structure and the kinetics of a reaction, an ODE model can be transcribed in SBML without reflecting the real structure of the reactions, hereby invalidating many structural analyses. We show that the automatic writing in SBML of the models of biomodels.net allows us to reduce the percentage of models with a non well-formed reaction from 66% to 28%.

6.2. Petri Net Analyses of Biochemical Networks using Constraint Logic Programming

Participants: François Fages, Thierry Martinez, Faten Nabli, Sylvain Soliman.

Petri nets are a simple formalism for modeling concurrent computation. Recently, they have emerged as a promising tool for modeling and analyzing biochemical interaction networks, bridging the gap between purely qualitative and quantitative models. Biological networks can indeed be large and complex, which makes their study difficult and computationally challenging.

In [10], we focus on two structural properties of Petri nets, siphons and traps, that bring us information about the persistence of some molecular species. We present a Boolean model and two constraint-based methods for enumerating all minimal siphons and traps of a Petri net, by iterating the resolution of Boolean satisfiability problems executed with either a SAT solver or a CLP(B) program. We compare the performances of these methods with respect to a state-of-the-art algorithm from the Petri net community. On a benchmark with 80 Petri nets from the Petriweb database and 403 Petri nets from curated biological models of the Biomodels database, we show that miniSAT and CLP(B) solvers are overall both faster by two orders of magnitude with respect to the dedicated algorithm. Furthermore, we analyse why these programs perform so well on even very large biological models and show a polynomial time complexity result for Petri nets of fixed treewidth, using a similar theorem for constraint satisfaction problems with bounded treewidth constraint graphs.

In [5] we present a method to compute the minimal semi-positive invariants of a Petri net representing a biological reaction system, as resolution of a Constraint Satisfaction Problem. This analysis brings both qualitative and quantitative information on the models, in the form of conservation laws, consistency checking, etc. thanks to finite domain constraint programming. It is noticeable that some of the most recent optimizations of standard invariant computation techniques in Petri nets correspond to well-known techniques in constraint solving, like symmetry-breaking. A simple implementation based on GNU-Prolog's finite domain solver, and including symmetry detection and breaking, was incorporated into the BIOCHAM modelling environment and in the independent tool Nicotine. Some illustrative examples and benchmarks are provided.

6.3. Subgraph Epimorphisms

Participants: François Fages, Steven Gay, Thierry Martinez, Francesco Santini, Sylvain Soliman.

The operations of deleting and merging vertices are natural operations for reducing a graph. While graph reductions through a sequence of vertex deletions (resp. mergings) characterize subgraph isomorphisms (resp. graph epimorphisms), sequences of both vertex deletion and merging operations characterize subgraph epimorphisms. Our proposal is thus to use subgraph epimorphism for comparing graphs in applications in systems biology and image analysis, when a more flexible notion than the classical notion of subgraph isomorphism is required.

In collaboration with Christine Solnon (INSA Lyon), we have developed the theory of subgraph epimorphisms. We have defined the SEPI, EPI and SISO distances between two graphs as the size of the largest SEPI (resp. EPI, SISO) lower bound graphs. These distances are equal to the minimum number of respectively vertex deletion and/or merging operations that are necessary to obtain isomorphic graphs. They are also metrics on graphs and we have $d_d \geq d_{md}$ and $d_m \geq d_{md}$. From a computational point of view, we have shown that the existence of a SEPI between two graphs is an NP-complete problem and have presented a constraint satisfaction algorithm for solving it.

Our algorithm is implemented in BIOCHAM and is currently improved for better performance on large graphs and generalized as a SEPI graph constraint propagation algorithm for computing SEPI lower and upper bounds.

6.4. Parameter Search with Temporal Logic Constraints

Participants: Grégory Batt, François Fages, Anthony Lins, Sylvain Soliman, Pauline Traynard, Jannis Uhlendorf, Luma Vittorino.

Our method for solving temporal logic constraints in first-order linear time logic LTL(R_{lin}), opens up the field of model-checking to optimization through the definition of a continuous degree of satisfaction for temporal logic formulae. This satisfaction degree can be used in a number of ways, e.g. as a fitness function with continuous optimization methods to find unknown parameter values in a model, to perform sensitivity analyses and compute the robustness of a system w.r.t. a temporal property and a perturbation of the parameters. or to find control parameters.

This approach is implemented in BIOCHAM and is one unique feature of this modeling environment. In this implementation, the continuous optimization procedure we use is the Covariance Matrix Adaptation Evolutionary Strategy CMAES of Nikolaus Hansen from the EPI TAO. A parallel version of Biocham implements this method on the Jade cluster of 10000 cores at GENCI for running our most challenging parameter search problems.

This year, in collaboration with Fernando Buarque, we have explored another continuous optimization method of the family of Particle Swarm Optimization (PSO), called Fish School Optimization (FSS). In [13], we report on our first results which are encouraging for using FSS for decreasing the sensitivity of the method to initial conditions and being able to maintain several swarms of solutions.

6.5. Coupled Model of the Cell Cycle and Circadian Clock

Participants: François Fages, Sylvain Soliman, Denis Thieffry, Pauline Traynard.

Recent advances in cancer chronotherapy techniques support the evidence that there exist imortant links between the cell cycle and the circadian clock genes. One purpose for modeling these links is to better understand how to efficiently target malignant cells depending on the phase of the day and patient characterictics. This is at the heart of our participation in collaboration with the EPI BANG in the EraNet SysBio project C5Sys, follow up of the former EU STREP project TEMPO.

This year we have investigated the effect of transcription inhibition during mitosis, as a reverse coupling from the cell cycle to the circadian clock. We use temporal logic constraints and the parallel version of BIOCHAM for parameter search, running on the Jade cluster of 10000 processors at the GENCI CINES, to couple dynamical models in high dimension and fit models to experimental data time series obtained in Franck Delaunay's lab in Nice, CNRS.

6.6. STL-based Analysis of TRAIL-induced Apoptosis

Participants: Grégory Batt, François Bertaux, Szymon Stoma.

Extrinsic apoptosis is a programmed cell death triggered by external ligands, such as the TNF-related apoptosis inducing ligand (TRAIL). Depending on the cell line, the specific molecular mechanisms leading to cell death may significantly differ. Precise characterization of these differences is crucial for understanding and exploiting extrinsic apoptosis. Cells show distinct behaviors on several aspects of apoptosis, including (i) the relative order of caspases activation, (ii) the necessity of Mitochondria Outer Membrane Permeabilization (MOMP) for effector caspase activation, and (iii) the survival of cell lines overexpressing Bcl2, leading to classification of cell lines into two groups (type I and type II). In [21], we challenge this type I/II cell line classification. We encode the three aforementioned distinguishing behaviors in a formal language, called signal temporal logic (STL), and use it to extensively test the validity of a previously-proposed model of TRAIL-induced apoptosis with respect to experimental observations made on different cell lines. Then, STL-guided parameter search is used to solve the few inconsistencies found between model and data. We show that these three criteria do not define consistent cell line classifications in type I or type II, and suggest mutants that are predicted to exhibit ambivalent behaviors. In particular, this finding sheds light on the role of a feedback loop between caspases, and reconciliates two apparently-conflicting views regarding the importance of either upstream or downstream processes for cell type determination. More generally, our work suggests that rather than being considered as defining criteria for cell type classification, these three distinguishing behaviors should be merely considered as type I or II features. On the methodological point of view, this work illustrates the biological relevance of STL-diagrams, STL population data, and STL-guided parameter search. Such tools are well adapted to the ever-increasing availability of heterogeneous knowledge on complex signal transduction pathways.

6.7. Real-time Control of Gene Expression in Yeast

Participants: Grégory Batt, François Fages, Jannis Uhlendorf, Jean-Baptiste Lugagne, Artémis Llamosi, Pascal Hersen.

Gene expression plays a central role in the orchestration of cellular processes. The use of inducible promoters to change the expression level of a gene from its physiological level has significantly contributed to the understanding of the functioning of regulatory networks. However, from a quantitative point of view, their use is limited to short-term, population-scale studies to average out cell-to-cell variability and gene expression noise and limit the nonpredictable effects of internal feedback loops that may antagonize the inducer action. In this project, in collaboration with the Hersen Lab at MSC (Paris Diderot University), we show that, by implementing an external feedback loop, one can tightly control the expression of a gene over many cell generations with quantitative accuracy. To reach this goal, we developed a platform for real-time, closed-loop control of gene expression in yeast that integrates microscopy for monitoring gene expression at the cell level, microfluidics to manipulate the cells environment, and original software for automated imaging, quantification, and model predictive control. By using an endogenous osmostress responsive promoter and playing with the osmolarity of the cells environment, we show that long-term control can, indeed, be achieved for both time-constant and time-varying target profiles at the population and even the single-cell levels [6]. Importantly, we provide evidence that real-time control can dynamically limit the effects of gene expression stochasticity. We anticipate that our method will be useful to quantitatively probe the dynamic properties of cellular processes and drive complex, synthetically engineered networks.

6.8. Genome Engineering of Mammalian Cells: Targeted and Efficient Integration of Multi-unit Genetic Payloads

Participants: Grégory Batt, Xavier Duportet.

Targeted integration of multi-unit genetic payloads would greatly benefit elucidating complex cellular mechanisms and implementing new functions in mammalian cells. Current technologies are however time-consuming and require tedious post-integration controls. To address this problem, we propose a modular framework to assemble large multi-unit genetic payloads and target their integration into either one or both alleles of a chromosomal locus of choice. To achieve this, we combine in a two-step process the customizable targeting properties of homing endonucleases with the efficiency and specificity of a large serine recombinase. We have demonstrated that an optimized version of BxB1 recombinase allows the targeted integration of large genetic circuits (up to 7 transcription units, 60kb) into a preintegrated landing pad in the AAVS1 locus, with a significant increase in efficiency compared to other site-specific recombination systems (integration in 10% of transfected cells without selection) . By reducing the time and efforts to generate large populations of isogenic stable cell lines adapted to study multi-component genetic systems, our framework is a valuable tool for mammalian synthetic biology and offers great potential for a broad range of biotechnology and therapeutic applications.

6.9. Reifying Global Constraints

Participants: François Fages, Raphaël Martin, Thierry Martinez, Sylvain Soliman.

Global constraints were introduced two decades ago as a means to model some core aspects of combinatorial problems with one single constraint for which an efficient domain filtering algorithm can be provided, possibly using a complete change of representation. However, global constraints are just constraint schemas on which one would like to apply usual constraint operations such as reification, i.e. checking entailment, disentailment and negating the constraint. This is currently not the case in state-of-the-art tools and was not considered in the global constraint catalog until recently. In [20], we propose a general framework for reifying global constraints and apply it to some important constraints of the catalog, such as the cumulative constraint for instance. We show that several global constraints that were believed to be hard to negate can in fact be efficiently negated, and that entailment and disentailment can be efficiently tested. We also point out some new global constraints that are worth studying from this point of view and provide some performance figures obtained with an implementation in Choco.

This scheme is currently used for compiling the Rules2CP constraint modeling language to Choco, and to internalize search in CSPs through constraint reification.

6.10. Railway Time Tabling Optimization

Participants: François Fages, David Fournier, Thierry Martinez, Sylvain Soliman.

Metros are able to generate electricity on a metro line by braking. This energy is immediately available in the third rail and is lost if no metro in the neighbourhood can consume it. It is thus possible to decrease the total energy consumption of a metro line by synchronizing the accelerations and braking of the metros. In [2], [9], we propose a classification of energy optimization timetable problems and we present a model for optimizing energy consumption which does not significantly alter the quality of service, by subtly modifying dwell times. We show however that this optimization problem is NP-hard. We present a hybrid genetic/linear programming algorithm for computing the distribution of braking metros. In this hybridization, the objective function is computed by a linear program and by a heuristic, and the dwell times are modified by a genetic algorithm. On a typical example with real data, the savings exceed 7%. Furthermore, on a benchmark of the literature for a simpler problem, we discuss the results obtained with our genetic algorithm, a tabu search algorithm and the mixed integer linear program used by the authors.

DEDUCTEAM Team

5. New Results

5.1. Dedukti

Together with Mathieu Boespflug (McGill University), Quentin Carbonneaux and Ronan Saillard have developed a new version of the front-end of Dedukti, written in OCaml, replacing an inefficient previous version, as well as a new version of the back-end using the Lua Just-In-Time compiler.

Ronan Saillard has internalized the Lua back-end of Dedukti, so that it is no longer necessary to explicitly call it when using Dedukti.

Ronan Saillard has extended the input language of Dedukti to allow the user to declare dependencies between modules, to write definition or to explicitly require to type-check a term.

Ronan Saillard has added a new feature to Dedukti to make opaque definitions. As with usual definitions, the proof term of an opaque definition is type-checked, but it is then immediately forgotten in order to decrease memory consumption.

5.2. Embeddings in the $\lambda\Pi$-calculus modulo

Ali Assaf has designed an embedding of the HOL logic in the $\lambda\Pi$-calculus modulo and implemented it in the HOLiDe system [40].

Together with Mathieu Boespflug, Ali Assaf and Guillaume Burel have developed an embedding of the Calculus of Inductive Constructions with universes in the $\lambda\Pi$-calculus modulo and Ali Assaf is currently implementing it in a new version of the CoqInE system.

Catherine Dubois and Raphaël Cauderlier have studied a translation in the $\lambda\Pi$-calculus modulo of features coming from object oriented programming languages, such as inheritance and late binding. This compilation scheme has been applied to produce a new back-end for FoCaLize [8], through a compilation to Dedukti. This new back-end is expected to be lighter than the present one producing Coq code and also to be able to combine local and external proofs coming from different proof environments [44]. They are currently working on a translation of the full FoCaLize language—not restricted to its object oriented features—and on a proof of its correctness with respect to the existing FoCaLize semantics.

5.3. Automated Theorem Proving

Guillaume Burel has shown that presenting theories by means of rewriting rules in Deduction modulo leads to more efficient proof search methods than using axioms, provided the rewriting system enjoys a proof theoretical property, namely cut admissibility.

He has been investigating which theories can be encoded as rewriting systems admitting cuts. Surprisingly, it turned out that any consistent theory in predicate logic can. This has been shown by studying the links between the set-of-support strategy of the Resolution method and the extension of the method based on Deduction modulo. He has also shown how to reduce the size of the corresponding rewriting systems [42].

Guillaume Burel has also studied how to improve the confidence in iProver Modulo. When it finds a resolution proof, it is now able to produce a proof that can be checked by Dedukti. The encoding of Resolution proofs in the $\lambda\Pi$-calculus modulo that is used is shallow, making more plausible the long-term goal of interoperability of provers, both interactive and automated, through Dedukti.

Simon Cruanes has explored several ideas for combining the Superposition calculus—one of the most powerful calculi for automated reasoning within first-order logic with equality—with Deduction modulo. Combining the term rewriting system for a theory in Deduction modulo with the ordered rewriting on which Superposition is based on proved to be difficult, yielding incomplete calculi; in most cases it boils down to the fact that the combination of confluent terminating term rewriting systems is in general neither terminating nor confluent. In order to experiment quickly ideas by implementing them, he has written a Superposition-based prover in OCaml, with some special features—automatic ordering of rewrite rules in the input, non-clausal calculus to be able to use equivalence relations as rewrite rules. The prover is 8,000 lines of code and is designed to be flexible and modular, but still has decent performance and can prove some non-trivial theorems.

Together with Mélanie Jacquel (Cedric), David Delahaye and Catherine Dubois have investigated Zenon for verifying proof rules added to help the automation in the provers of Atelier B. They have augmented Zenon with specific rules for dealing with set operations and predicates, obtained by applying super deduction—a variant of Deduction modulo [33].

5.4. Proof theory

We believe that our work on proof-checking and automated theorem proving cannot be separated from a more theoretical research on proof theory.

Together with Denis Cousineau, Gilles Dowek and Olivier Hermant have related semantic criteria for proof normalization and admissibility of the cut rule in Deduction modulo [17], [26].

Gilles Dowek has proposed a new way to define classical connectives in a constructive framework [46].

Together with Murdoch J. Gabbay (Heriot Watt), Gilles Dowek has proposed a new nominal logic that handles binders in terms [16] and a new semantics for predicate logic [29].

During her visit in the team, Cecilia Englander has studied the correspondence between natural deduction and sequent calculus.

Together with Ying Jiang (Beijing), Gilles Dowek has defined a logic for finite structures. Kailiang Ji is currently investigating the use of proof search algorithms in Deduction modulo to automatically prove theorems in this theory.

5.5. Safety of aerospace systems

Together with Anthony Narkawicz (Nasa-Langley) and César Muñoz (Nasa-Langley), Gilles Dowek has designed a prevention bands algorithm, that is an algorithm that computes and displays to the pilot of an aircraft, a sequence of safe and unsafe intervals on ground speed, heading or vertical speed and they have proved this algorithm correct in the PVS system [18].

This algorithm computes with real numbers, but its implementation computes with floating point numbers. Moreover this algorithm is numerically unstable as it uses comparisons of numbers, computed with square root and division operations. This has led Pierre Néron to design a program transformation algorithm to eliminate square roots and divisions in straight-line programs. This way computation can be made exact.

Together with César Muñoz, Pierre Néron has completed this year the design of this program transformation algorithm and he has proved, in the PVS system, its termination and correctness: preservation of semantics and absence of square roots and divisions in the produced program [35].

Together with César Muñoz, Pierre Néron has also implemented this transformation algorithm as a PVS automatic proof strategy, that allows a wider range of expressions, using a deep embedding of PVS in PVS itself.

Pierre Néron and Raphaël Bost have proposed an optimization of one aspect of that algorithm: the definition of a common template for arithmetic expression.

5.6. Constraint Solving

Catherine Dubois has developed in collaboration with Matthieu Carlier and Arnaud Gotlieb (Oslo) a formally verified constraint finite domain solver. It focuses on arc-consistency and has been developed with Coq [24].

5.7. Models of Computation

Together with Pablo Arrighi (Grenoble), Gilles Dowek has reformulated Gandy's proof of the physical Church-Thesis in the quantum case [11]. Gilles Dowek has proposed the idea that the Galileo thesis could be seen as a consequence of the physical Church-Turing thesis and therefore as a consequence of Gandy's principles [15]. Gilles Dowek has proposed a definition of a notion of non deterministic computation over the real numbers [14] that could be used as a language to describe continuous non deterministic physical phenomena. All this work has then been presented in a tutorial at the conference *Language and Automata Theory and Applications* [28].

Together with Pablo Arrighi, Gilles Dowek has investigated further the principle of a finite density of information [38] and in particular the impact of this definition on the notion of a chaotic dynamical system [37].

Together with Pablo Arrighi, Gilles Dowek has investigated a generalization of the notion of cellular automaton where the principle of a bounded density of information is formulated independently of the geometry of space. This led to the notion of a Causal graph dynamic [12].

Nachum Dershowitz and Gilles Dowek have shown that extending Turing machines with a two-dimensional tape, made this formalism usable in practice to implement classical algorithms [45].

Alejandro Díaz-Caro and Gilles Dowek have proposed to take a fresh look at non deterministic λ-calculi—such as quantum λ-calculi—and derive non determinism from type isomorphism [30].

Together with Giulio Manzonetto (Paris 13) and Michele Pagani (Paris 13), Alejandro Díaz-Caro has considered an extension of the call-by-value λ-calculus with a may-convergent non-deterministic choice and a must-convergent parallel composition, endowed with a type system. They have proved that a term is typable if and only if it is converging, and that its typing tree carries enough information to give a bound on the length of its lazy call-by-value reduction. Moreover, when the typing tree is minimal, such a bound becomes the exact length of the reduction [31].

Together with Barbara Petit (Sardes), Alejandro Díaz-Caro has considered the non-deterministic extension of the call-by-value lambda calculus, which corresponds to the additive fragment of the linear-algebraic lambda-calculus. They have defined a fine-grained type system, capturing the right linearity present in such formalisms. After proving the subject reduction and the strong normalisation properties, they have proposed a translation of this calculus into the System F with pairs, which corresponds to a non linear fragment of linear logic. The translation provides a deeper understanding of the linearity in this setting [32].

Together with Pablo Arrighi, Barbara Petit, Pablo Burias (Rosario), Mauro Jaskelioff (Rosario), and Benoît Valiron (Penn), Alejandro Díaz-Caro has studied possible typing systems for the full linear-algebraic λ-calculus in which the non-deterministic calculus can be seen as a particular case. They have proposed a type system that keeps track of "the amount of a type" that is present in each term [13]. As an example of its use, they have shown that it can serve as a guarantee that the normal form of a term is barycentric, that is that its scalars are summing to one. They also proposed a type system similar to the one presented in [32], but for the full calculus, ensuring confluence and convergence [23]. Finally, they provided a full type system that is able to statically describe the linear combinations of terms resulting from the reduction of programs, also ensuring convergence [19].

FORMES Team

6. New Results

6.1. Higher-Order Abstract Syntax

This recently started project funded by the National Science Foundation of China aims at setting up a generic infrastructure for representating logical systems and automate their meta-theoretical study. We view a logical system as a type theory made of three components: a language of terms, types being particular terms; a set of typing rules; and a set of computational rules described by typed higher-order rewrite rules.

There are several challenges in this project. The first is to define logical frameworks which are expressive enough -at least as expressive as Girard's System F or Edingburgh's LF- to define the syntax and semantics of rich type theories, such as CoqMTU as an extreme example. A second challenge is to develop new techniques for checking the three main properties of higher-order rewrite rules: type preservation -which is usually easy-, confluence and termination. Our work here has progressed steadily, in paticular with new advanced techniques for checking termination and confluence described next. A third challenge is to formalize these results in Coq, in order to provide proof certificates for particular cases. The fourth challenge is to build a a general infrastructure in Coq in which all these techniques become available in order to study particular logical systems.

As initial steps, we undertook the following formalizations :

- Hua Mei implemented an intensional framework for simply typed lambda-calculus in Coq, where α- and β-conversions have been axiomized.
- Frédéric Blanqui has formalized in Coq the pure lambda-calculus following the definition of Curry and Feys in [43] (named variables and explicit alpha-equivalence), and the proof of termination of β-reduction for simply-typed λ-terms based on computability predicates [51]. To the best of his knowledge, this is the first formalization of the termination of β-reduction using named variables and explicit alpha-equivalence, all the other formalizations using De Bruijn indices [73] or nominal logic [48].
- Qian Wang formalized completely the theory of CoqMTU in Coq augmented with strong set-theoretic axioms in order to get around Gödel's incompleteness theorem. This is described in more details next.

6.2. CoqMTU

The proof-assistant Coq is based on a complex type theory, which resulted from various extensions of the Calculus of Constructions studied independently fromf each other. With Bruno Barras, we decided to address the challenge of proving the real type theory underlying Coq, and even, indeed, its recent extension CoqMT. To this end, we have studied formally the theory CoqMTU, which extends the calculus of Constructions with inductive types, a predicative hierarchy of universes and a decidable theory T for some first-order inductive types for which large elimination is no more available. This work has been published at LICS [1]. It leaves open the question whether large elimination can be accomadated for those inductive types which carry along a decidable theory T. This problem has been solved recently by Wang, who constructed a set-theoretic model of CoqMTU with strong elimination.

6.3. Normal Rewriting

There are many forms of rewriting used in the litterature: plain rewriting (rules are fired via plain pattern matching), rewriting modulo T (rules are fired via pattern matching modulo T), higher-order rewriting (rules are fired via higher-order pattern matching, but apply to simply typed lambda-terms terms provided the redex is of base type and in beta-normal eta-long form). For each of these rewriting mechanisms, there are results describing how to check confluence and termination.

Regarding confluence, these results describe which *critical pairs* must be computed in order to check the confluence property of the rewriting relation, assuming some termination property. In [17], we describe a general abstract result which can then be instantiated to all of the previous cases, and removes the assumptions above for higher-order rewriting. This is done via two novel notions: abstract positional rewriting allows us to capture the notion of critical peak without having to talk about a specific term structure; abstract normal rewriting with a triple (R, S, E) allows us to capture all different forms of rewriting: $S = E = \varnothing$ for plain rewriting; $S = \varnothing$ for rewriting modulo; E is alpha-conversion for higher-order rewriting, while the set of simplifiers S is made of beta-reduction and eta-expansion, R being the set of user-defined rules. Of course, there are other applications of normal rewriting described in the paper: for first-order computations, but also for higher-order computations at higher types, or using eta-reduction instead of eta-expansion, therefore solving a long-standing open problem.

Regarding termination, these results are very preliminary. In a recent paper submitted to ACM Transactions on Computational Logics, we extend the termination proof methods for higher-order computations based on plain pattern matching to higher-order rewriting systems based on higher-order pattern matching. We accomodate, for the one hand, with a weakly polymorphic, algebraic extension of Church's simply typed λ-calculus, and on the other hand, with any use of eta, as a reduction, as an expansion or as an equation. User's rules may be of any type in this type system, either a base, functional, or polymorphic type. Our techniques fit well with higher-order reduction orderings, such as the computability path ordering, but can also be used by other techniques, such as higher-order dependency pairs. All examples of normal higher-order rewrite rules that can be found in the litterature can be treated by our techniques, even those for which termination is by no means obvious to the expert.

6.4. Decreasing Diagrams

Based on the so-called Newman's lemma, the method for checking confluence introduced in the former paragraph applies to terminating computations. A completely different technique based on the so-called Hindley-Rosen's lemma applies when computation do not terminate, and is at the basis of Tait's confluence proof for the pure lambda-calculus. In recent papers, van Oostrom succeeded to capture both within a single framework thanks to the notion of decreasing diagram of a labelled abstract relation [76], see also [11] for an improved proof. Decreasing diagrams are specific convertibility proofs for local peaks, which labels are smaller in some sense than those of the local peak they aim at replacing. Any convertibility proof can then be converted into a confluence proof by recursively replacing its local peaks by their associated decreasing diagrams. Using a subtle characterization of confluence for arbitrary (possibly non-terminating) relations by cofinal derivations due to Klop [11], van Oostrom showed that any confluent relation which convertibility classes are countable, can be labelled in a way that makes it a labelled relation satisfying the decreasing diagram condition.

In [15], we first give a new, simple proof of van Oostrom's initial result based on a subtle well-founded order on conversions, and generalize it to rewriting modulo by using *strongly coherent cliffs* as an analog of decreasing diagrams for peaks. We then extend Klop's cofinal derivations to *cofinal streams*, and prove again a completeness result under the strong coherence assumption. Finally, we derive from these results a new, compact proof of Toyama's theorem that confluence is a modular property of rewriting systems built on disjoint vocabularies, and extend it to rewriting modulo when strong coherence is satisfied.

We are now trying to get rid of the strong coherence assumption by introducing a weaker analog of decreasing diagrams, *decreasing cliffs*. A preliminary result was presented early november at the Japanese Term Rewriting Workshop in Sendai.

This line of work is very promising. We expect it will eventually lead to the solution of an old open problem, the characterization of a class of non-left linear, non-terminating rewrite systems for which confluence is decidable by means of (parallel) critical pairs. We believe that the implementation of such a result would be impact the way confluence proofs are carried out, including in type theory.

6.5. Higher-order Reduction Orderings

Since HORPO , several higher-order reduction orderings have been described, based on either Dershowitz's RPO , Blanqui-Jouannaud-Okada's Computational Closure , and Arts and Giesel' dependency pairs . Our work continues in three different directions:

- CPO is an order for simply typed lambda-terms that allows to show strong normalization of beta-reduction even in presence of higher-order rewrite rules provided these rules decrease in the ordering [32]. It is currently the only automated mechanism that achieves non-trivial computations by turning Girard's computability predicates method into a usable tool. It has been shown that CPO can handle weakly polymorphic type disciplines, as well as inductive types. Recently, we have shown that CPO scales up to dependently typed calculi as LF. We are currently writing a paper describing CPO and its extensions to calculi with inductive and dependent types which should be submitted to a journal by the end of the year.
- Frédéric Blanqui defended his "Habilitation à diriger des recherches" at the University Denis Diderot (Paris 7) on July 13. In [13], he gives a synthetic view on how the notion of computability closure can be used to prove the termination of various kinds of rewrite relations (class rewriting or rewriting with matching modulo), and how it relates with other notions (dependency pairs, semantic labeling, and HORPO, the predecessor of CPO.
- Frédéric Blanqui has developed an automated termination prover called HOT based on the above work on the computability closure and his former work on size annotations [31]. For its first participation, HOT won the international competition on termination in the category "higher-order rewriting union beta".

6.6. Certification of Termination Proofs

Frédéric Blanqui and Kim Quyen Ly continued to work on the development of a new version of Rainbow based on Coq extraction mechanism [59]. We developed a tool generating from an XSD file, Coq and OCaml data structures representing the XML types defined the XSD file, and OCaml parsing functions for generating such data structures from an XML file. The main difficulty was to topologically reorder the XSD type definitions in order to get simple and well defined Coq data structures. We also defined and proved in Coq a function for checking the correctness of termination certificates based on the DP transformation [26]. The main difficulty was to manage the evolution of the arity function along the transformation. Indeed, to simplify the translation of CPF elements into the data structures used in CoLoR [30], we decided to use a fixed but infinite set of symbols [69]. However the arity function need to be updated along the transformations applied to the system. These results are presented in [20].

6.7. Certification of Moca

Frédéric Blanqui has formalized in Coq and proved the correctness and completeness of the construction functions generated by Moca for the theory of groups [29]. The first difficulty is to represent the Moca functions themselves in a faithful way because, in Coq, there is no "when" clauses and "match" constructions are expanded into elementary "case" constructions with no tuple patterns and patterns of depth one only. In addition, Coq termination checker only accepts functions with exactly one structurally decreasing argument, which is generally not the case of Moca functions. The second difficulty is the completeness proof: it requires the use of intermediate data structures for reasoning on normal forms. During his internship, Rémi Nollet (L3, ENS Lyon) improved the representation of OCaml functions by using inductive predicates, and extended the correctness proof to commutative groups.

6.8. First steps towards the certification of an ARM simulator

The simulation of Systems-on-Chip (SoC) is nowadays a hot topic because, beyond providing many debugging facilities, it allows the development of dedicated software before the hardware is available. Low-consumption

CPUs such as ARM play a central role in SoC. However, the effectiveness of simulation depends on the faithfulness of the simulator. To this effect, we started to prove significant parts of such a simulator, SimSoC. Basically, on one hand, we develop a Coq formal model of the ARM architecture while on the other hand, we consider a version of the simulator including components written in Compcert-C [58]. Then we prove that the simulation of ARM operations, according to Compcert-C formal semantics, conforms to the expected formal model of ARM. Size issues are partly dealt with using automatic generation of significant parts of the Coq model and of SimSoC from the official textual definition of ARM [3]. A second step was achieved in [12], with the proof a significant instruction (ADC, Add with Carry). A crucial technical issue was then raised: facilitating reasoning by inversion on the rules defined in Compcert-C. Hundreds such steps are required for a single instruction, and each of them generates a dozen of new names. Relying on Coq tactic inversion results in unmanageable scripts, very fragile and difficult to maintain. In 2012 we dealt with this issue by designing our own inversion mechanism, allowing us to improve automation of the proof, while keeping enough command so that interactive steps refer to controlled names. It was then possible to get a much shorter proof on ADC and to prove at least one instruction in each category of the ARM instruction set.

6.9. Certified implementation of BIP

BIP (*Behavior, Interaction, Priority*) is a component-based language designed at VERIMAG for modeling and programming complex embedded systems [27]. A BIP model is essentially a set of atomic components described with explicit states and transitions, composed together in a hierarchical way. The main original feature of BIP lies in a very rich notion of *connector* for defining interactions between components [33]. An efficient implementation of BIP in C++ is already available at VERIMAG.

Building on our previous experience on SimSoC, we started to work on a certified implementation of BIP. Our long term objective is to propose a certified compilation chain from BIP models to embedded code, through a first translation from BIP to Compcert-C.

In 2012 we focused on a simple subset of BIP Currently, we have a first definition of a formal semantics of this subset in Coq, in two versions: an relational version, inspired by a rule-based operational semantics, and a functional version, which specifies a possible implementation of the relational version (in particular, it includes a scheduler). We also produce a Compcert-C code which is expected to behave exactly like the functional semantics, and we started to state and prove corresponding statements on very simple BIP models.

6.10. Formal model and proofs for Netlog protocols

Netlog is a language designed and implemented in the Netquest project for describing protocols. Netlog has a precise semantics, provides a high level of abstraction thanks to its Datalog flavor and benefits from an efficient implementation. This makes it a very interesting target language for proofs of protocols.

Jean-François Monin, Stéphane Grumbach (formerly LIAMA/Netquest) and Yuxin Deng (Jiaotong University, Shanghai) designed a formal model of Netlog in Coq, where the two possible semantics are derived from common basic blocks. In a fully certified framework, a formal proof of the Netlog engine (running on each node) would be required. We don't attack this part at the moment: we assume that the implementation respects the general properties stated in our model and focus on the issues raised by the distributed model of computation provided by Netlog. This framework could be applied to an algorithm constructing a Breadth-First Search Spanning Tree (BFS) in a distributed system [45].

In 2011, Jean-François Monin and Meixian Chen (Jiaotong Shanghai) generalized the model in order to take the removal of datalog facts into account, and used the improved framework to Prim's algorithm. In 2012, this work was slightly improved and published in [16].

6.11. Formalisation of security APIs for mobile phones

This work is in cooperation with Nokia Beijing, who was interested by the application of verification technologies to mobile phones. We decided to focus on security APIs, considering that mobile devices are commonly used by end-users to store their personnal data (e.g., passwords), while running all sort of downloaded applications at the same time.

For 2012, we (including Nokia) agreed to consider devices under Android, though Nokia switched to windows, in order to circumvent copyright issues.

Three models and corresponding sets of APIs for password storage applications on Android were developed. Each model fixes some bugs of the previous one and introduces a new feature. We consider the third model is enough for the basic function and well built to be safe. Then, a full Coq proof of the third model was developed as well as its corresponding API's security property. A suitable abstraction of the application on the phone within its environment is described as a state transition system. Then we proved by induction that the expected secrets actually remain secret at any reachable state.

6.12. Trace Analysis

Simulation sessions produce huge trace files, sometimes now in hundreds of gigabytes, that are hard to analyze with a quick response time. This comes down to two sub-problems:

- The trace file size. Trace files are huge because they include lots of information. But when looking for a specific problem, one does not need all of this information. To search one given defect, one may ignore a large amount of the data in the trace file. One would like the trace file to contain only relevant information to the concerned problem.
- The expressive power of the language to analyze the trace, and its usability. If the language is limited to expression search, it is easy to use but hard to construct sophisticated formulas. If the language used is Linear Temporal Logic (LTL), there is a very high expressive power but many engineers are unable to write a LTL formula and to maintain it over time.

We have started to build a trace analysis tool. It includes a language which allows expression of time-related formulas as a subset of LTL, but is simple to formulate expressions. When this language is compiled, the compiler generates two outputs:

- a filter script that will help reduce the size of the trace file.
- a program that analyzes such trace files to find whether the formula is satisfied.

When compiling one trace language input file, it generates a filter script. The filter script is a set of data descriptors. It describes which events from the simulator must be traced and which should be ignored. Then during the simulation, the filter is loaded and only the required output is generated.

We have started to design a trace language and a compiler, and extended the SimSoC simulator to support generation of trace files with a filter. A first version of the trace language compiler has been implemented in OCAML, which generates OCAML programs for trace analysis. In the current version under development, the filters are not yet parallelized with simulation.

GALLIUM Project-Team

6. New Results

6.1. Language design and type systems

6.1.1. The Mezzo programming language

Participants: Jonathan Protzenko, François Pottier.

In the past ten years, the type systems community and the separation logic community, among others, have developed highly expressive formalisms for describing ownership policies and controlling side effects in imperative programming languages. In spite of this extensive knowledge, it remains very difficult to come up with a programming language design that is simple, effective (it actually controls side effects!) and expressive (it does not force programmers to alter the design of their data structures and algorithms).

The Mezzo programming language, formerly known as HaMLet, aims to bring new answers to these questions.

We have come up with a solid design for the programming language: many features of the language have been reworked or consolidated this year, and we believe we strike a good balance between expressiveness and complexity. We wrote several flagship examples that illustrate the gains offered by Mezzo, as well as two (yet unpublished) papers discussing the design of the language. Jonathan Protzenko implemented a prototype type-checker; although it is not perfect yet, several non-trivial examples are successfully type-checked.

The current state of the Mezzo programming language is best described in [40]; a former version of this document can be found as [39].

François Pottier wrote a formal definition of (a slightly lower-level variant of) Mezzo, and proved that Mezzo is type-safe: that is, well-typed programs cannot crash (but they can stop abruptly if a run-time check fails). The proof, which is about 15,000 lines, has been machine-checked using Coq. A paper that describes this work is in preparation.

This work was facilitated by Pottier's experience with a similar previous proof. In particular, out of the above 15,000 lines, about 2,000 lines correspond to a re-usable library for working with de Bruijn indices, and about 3,000 lines correspond to a re-usable formalisation of "monotonic separation algebras", which help reason about resources (memory, time, knowledge, ...) and how they evolve over time. These libraries have not yet been fully documented and released; this might be done in the future.

6.1.2. Coercion abstraction

Participants: Julien Cretin, Didier Rémy.

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 studies of such systems is to make type conversions explicit as "coercions" inside terms.

Following a general approach to coercions based on System F, we introduced a language F-iota with abstraction over coercions and where all type transformations are represented as coercions. The main difficulty is dealing with coercion abstraction, as abstract coercions whose types are uninhabited cannot be erased at run-time. We proposed a restriction, called parametric F-iota, that ensures erasability of all coercions by construction. This work was presented at the POPL conference in January [22].

We extended parametric F-iota with non-interleaved positive recursive types and with erasable isomorphisms. We generalized the presentation of the language viewing coercions as conversions between typings (pairs of a typing environment and a type) rather than between types. An extended version with full proofs will be submitted for journal publication.

We also studied a more liberal version of F-iota where coercion inhabitation is no more ensured by construction (which limits expressiveness), but instead by providing coercion witnesses in source terms. This extension requires pushing abstract coercions under redexes so that they do not block the reduction. As a consequence, coercions cannot be reified in System F, and we need a direct proof of termination of iota-reduction. We completed one such proof based on reducibility candidates.

6.1.3. Ambivalent types for principal type inference with GADTs

Participants: Jacques Garrigue [Nagoya University], Didier Rémy.

Type inference for Generalized Abstract Data Types (GADTs) is always a matter of compromise because it is inherently non monotone: assuming more specific types for GADTs may ensure more invariants, which in turn may result in more general types. Moreover, even when types of GADTs parameters are explicitly given, they introduce equalities between types, which makes them inter-convertible but with a limited scope. This may then creates an ambiguity when leaving the scope of the equation: which representative should be used for the equivalent forms? Ideally, one should use a type disjunction, but this is not allowed—for good reasons. Hence, to avoid arbitrary choices, these situations must be rejected, forcing the user to add more annotations to resolve ambiguities.

We proposed a new approach to type inference with GADTs. While some uses of equations are unavoidable and create real ambiguities, others are gratuitous and create artificial ambiguities, To distinguish between the two, we introduced *ambivalent types*: a way to trace types that have been obtained by an unavoidable use of an equation. We then redefined ambiguities so that only ambivalent types become ambiguous and should be rejected or resolved by a programmer annotation.

Interestingly, the solution is fully compatible with unification-based type inference algorithms used in ML dialects. The work was presented at the ML workshop [31] and implemented in the latest version 4.00 of OCaml.

6.1.4. GADTs and Subtyping

Participants: Gabriel Scherer, Didier Rémy.

Following the addition of GADTs to the OCaml language in version 4.00 released this year, we studied the theoretical underpinnings of variance subtyping for GADTs. The question is to decide which variances should be accepted for a GADT-style type declaration that includes type equality constraints in constructor types. This question exposes a new notion of decomposability and unexpected tensions in the design of a subtyping relation. Our formalization partially reuses earlier work by François Pottier and Vincent Simonet [54]. It was presented at the ML Workshop [33]. An extended version including full proofs is available as a technical report [38] and was submitted for presentation at a conference.

6.1.5. Singleton types for code inference

Participants: Gabriel Scherer, Didier Rémy.

Inspired by tangent aspects of the PhD work of Julien Cretin, we investigated the use of 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. The preliminary results seem encouraging, both on the theoretical side (identifying general situations for type-directed programming) and the practical side (mining existing OCaml code for usage situations).

6.1.6. Programming with names and binders

Participants: Nicolas Pouillard, François Pottier.

Following Nicolas Pouillard's Ph.D. defense in January 2012 [11], Nicolas Pouillard and François Pottier produced a unified presentation of Pouillard's approach to programming with abstract syntax, in the form of a paper that was published in the Journal of Functional Programming [16].

6.1.7. A type-and-capability calculus with hidden state

Participant: François Pottier.

During the year 2010, François Pottier developed a machine-checked proof of an expressive type-and-capability system, which can be used to type-check and prove properties of imperative ML programs. The proof is carried out in Coq and takes up roughly 20,000 lines of code. In the first half of 2011, François Pottier wrote a paper that describes the system and its proof in detail. This paper was published, after a revision, in 2012 [15].

6.2. Formal verification of compilers and static analyses

6.2.1. The CompCert verified C compiler

Participants: Xavier Leroy, Sandrine Blazy [project-team Celtique], Jacques-Henri Jourdan, Valentin Robert.

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.

The two major novelties of CompCert this year are described separately: verification of floating-point arithmetic (section 6.2.2) and a posteriori validation of assembly and linking (section 6.2.3). Other improvements to CompCert include:

- The meaning of "volatile" memory accesses is now fully specified in the semantics of the CompCert C source language. Their translation to built-in function invocations, previously part of the unverified pre-front-end part of CompCert, is now proved correct.
- CompCert C now natively supports assignment between composite types (structs or unions), passing composite types by value as function parameters, and other instances of using composites as r-values, with the exception of returning composites by value from a function.
- A new pass was added to the compiler to perform inlining of functions. Its correctness proof raised interesting challenges to properly relate the (widely different) call stacks of the program before and after inlining.
- The constant propagation optimization is now able to propagate the initial values of global variables declared `const`.
- The common subexpression elimination (CSE) optimization was improved so as to eliminate more redundant memory loads.

Two versions of the CompCert development were publicly released, integrating these improvements: versions 1.10 in March and 1.11 in July. We also wrote a 50-page user's manual [37] and a technical report on the CompCert memory model [35].

In parallel, we continued our collaboration with Jean Souyris, Ricardo Bedin França and Denis Favre-Felix at Airbus. They are conducting an experimental evaluation of CompCert's usability for avionics software, and studying the regulatory issues (DO-178 certification) surrounding the potential use of CompCert in this context. Preliminary results were presented at the 2012 Embedded Real-Time Software and Systems conference (ERTS'12) [29].

6.2.2. Formalization of floating-point arithmetic in Compcert

Participants: Sylvie Boldo [project-team Toccata], Jacques-Henri Jourdan, Xavier Leroy, Guillaume Melquiond [project-team Toccata].

The aim of this research theme was to formalize the semantics and compilation of floating-point arithmetic in the CompCert compiler. Prior to this work, floating-point arithmetic was axiomatized in the Coq proof of CompCert, then mapped to OCaml's floating-point operations during extraction. This approach was prone to errors and fails to formally guarantee conformance to the IEEE-754 standard for floating-point arithmetic.

To remedy this situation, Jacques-Henri Jourdan replaced this axiomatization by a fully-formal Coq development, building on the Coq formalization of IEEE-754 arithmetic provided by the Flocq library. Sylvie Boldo and Guillaume Melquiond, authors of Flocq, adapted their library to the needs of this development. The new formalization of floating-point arithmetic is used throughout CompCert: to give semantics to FP computations in the source, intermediate and target (assembly) languages; to perform correct compile-time FP evaluations during constant propagation; to prove the correctness of code generation scheme for conversions between integers and FP numbers; and to parse FP literals with correct rounding.

A paper describing this work is accepted for presentation at the forthcoming ARITH 2013 conference [20].

6.2.3. Validation of assembly and linking

Participants: Valentin Robert, Xavier Leroy.

Valentin Robert designed and implemented a validation tool for the assembly and linking phases of the CompCert C compiler. These passes are not formally verified and call into off-the-shelf assemblers and linkers. The `cchecklink` tool of Valentin Robert improves the confidence that end-users can have in these passes by validating *a posteriori* their operation. The tool takes as inputs the PowerPC/ELF executable produced by the linker, as well as the abstract syntax trees for assembly files produced by the formally-verified part of CompCert. It then proceeds to establish a correspondence between the two sets of inputs, via a thorough structural analysis on the ELF executable, light disassembling of the machine code, expansion of CompCert's macro-asm instructions, and propagation of constraints over symbolic names. The tool produces detailed diagnostics if any discrepancies are found.

6.2.4. Improving CompCert's reusability for verification tools

Participants: Xavier Leroy, Jacques-Henri Jourdan, Andrew Appel [Princeton University], Sandrine Blazy [project-team Celtique], David Pichardie [project-team Celtique].

Several ongoing projects focus on proving the soundness of verification tools that reuse parts of the CompCert development, namely some of the intermediate languages, their formal semantics, and the CompCert passes that produce these intermediate forms. This is the case for the Verasco ANR project, which focuses on the proof of a static analyzer based on abstract interpretation, and for the Verified Software Toolchain (VST) project, led by Andrew Appel at Princeton University, which develops a concurrent separation logic embedded in Coq. However, the CompCert intermediate languages, currently designed to fit the needs of a compiler, are not perfectly suited to static analysis and deductive verification.

To improve the reusability of CompCert's Clight language in the Verasco and VST projects, Xavier Leroy is currently revising the CompCert C front-end passes so that function-local C variables whose address is never taken are pulled out of memory and replaced by nonadressable temporary variables. The resulting Clight intermediate form is much easier to analyze or prove correct, as temporary variables cannot suffer from aliasing problems.

Likewise, Sandrine Blazy, Jacques-Henri Jourdan, Xavier Leroy and David Pichardie designed a variant of CompCert's RTL intermediate language, called CFG. Like RTL, CFG represents the flow of control by a graph; unlike RTL, CFG is independent of the target processor, and supports complex expressions instead of 3-address code. These features of CFG make it a better target for static analysis, both non-relational (e.g. David Pichardie's certified interval analysis) and relational. Jacques-Henri Jourdan implemented and proved correct a compilation pass that produces CFG code from the Cminor intermediate language of CompCert.

6.2.5. Formal verification of hardware synthesis

Participants: Thomas Braibant, Adam Chlipala [MIT].

Verification of hardware designs has been thoroughly investigated, and yet, obtaining provably correct hardware of significant complexity is usually considered challenging and time-consuming. Hardware synthesis aims to raise the level of description of circuits, reducing the effort necessary to produce them.

This yields two opportunities for formal verification: a first option is to verify (part of) the hardware compiler; a second option is to study to what extent these higher-level design are amenable to formal proof.

During a visit at MIT, Thomas Braibant worked on the implementation and proof of correctness of a prototype hardware compiler in Coq, under Adam Chlipala's supervision. This compiler produces descriptions of circuits in RTL style from a high-level description language inspired by BlueSpec. After joining Gallium, Thomas Braibant continued working part time on this subject, finishing the proof of this compiler, and implementing a few hardware designs of mild complexity. This work was presented at the 2012 Coq Workshop [30] and will be submitted to a conference in 2013.

6.2.6. A formally-verified alias analysis

Participants: Valentin Robert, Xavier Leroy.

Valentin Robert improved the verified static analysis for pointers and non-aliasing that he initiated in 2011 during his Master's internship supervised by Xavier Leroy. This alias analysis is intraprocedural and flow-sensitive, and follows the "points-to" approach of Andersen [41]. An originality of this alias analysis is that it is conducted over the RTL intermediate language of the CompCert compiler: since RTL is essentially untyped, the traditional approaches to field sensitivity do not apply, and are replaced by a simple but effective tracking of the numerical offsets of pointers with respect to their base memory blocks. The soundness of this alias analysis is proved against the operational semantics of RTL using the Coq proof assistant and techniques inspired from abstract interpretation. A paper describing the analysis and its soundness proof was presented at the CPP 2012 conference [26].

6.3. The OCaml language and system

6.3.1. The OCaml system

Participants: Xavier Clerc [team SED], Damien Doligez, Alain Frisch [Lexifi SAS], Jacques Garrigue [University of Nagoya], Fabrice Le Fessant [Inria Saclay and OCamlPro start-up company], Jacques Le Normand [Lexifi SAS], Xavier Leroy.

This year, we released versions 4.00.0 and 4.00.1 of the OCaml system. Version 4.00.0 (released in July) is a major release that fixes about 150 reported bugs and 4 unreported bugs, and introduces 57 new features suggested by users. Version 4.00.1 (released in October) is a bug-fix release that fixes 3 major and 20 minor bugs. Damien Doligez acted as release manager for both versions.

The major innovation in OCaml 4.00 is support for generalized algebraic datatypes (GADTs). These non-uniform datatype definitions enable programmers to express some invariants over data structures, and the OCaml type-checker to enforce these invariants. They also support interesting ways of reflecting types into run-time values. GADTs are found in proof assistants such as Coq and in functional languages such as Agda and Haskell. Their integration in OCaml raised delicate issues of partial type inference and principality of inferred types, to which Jacques Garrigue and Jacques Le Normand provided original solutions [45].

Other features of this release include:

- Lightweight notations to facilitate the use of first-class modules.
- Better reporting of type errors.
- Improvements in native-code generation.
- Performance and security improvements in the hashing primitive and hash tables.
- New warnings for unused code (variables, record fields, etc.)
- A new back-end for the ARM architecture.

6.3.2. Namespaces for OCaml

Participants: Gabriel Scherer, Didier Rémy, Fabrice Le Fessant [Inria Saclay].

As part of an ongoing discussion among members of the OCaml Consortium, we investigated the formal aspects of "namespaces" and their putative status in the OCaml language. Namespaces aim at providing OCaml programmers with efficient ways to manage and structure the names of compilation units, in contrast with the flat, global space of compilation units provided today in OCaml. This formalization provides scientific support to ongoing design and engineering discussions. It was presented at the December 2011 IFIP 2.8 working group on functional programming, and at the December 2012 meeting of the OCaml Consortium.

6.4. Software specification and verification

6.4.1. Tools for TLA+

Participants: Damien Doligez, 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 [47], and to build tools for writing TLA+ specifications and mechanically checking the corresponding formal proofs.

This year, the TLA+ project released two new versions (in January and in November) of the TLA+ tools: the GUI-based TLA Toolbox and the TLA+ Proof System, an environment for writing and checking TLA+ proofs. This environment is described in a paper presented at the 2012 symposium on Formal Methods [21]. The January release (version 1.0 of TLAPS and 1.4.1 of Toolbox) added support for back-ends based on SMT provers (CVC3, Z3, Yices, VeriT), which dramatically extends the range of proof obligations that the system can discharge automatically. The November release includes many bug-fixes and performance improvements.

We have also improved the theoretical design of the proof language with respect to temporal properties. This design will be implemented in TLAPS in the near future.

Web site: http://tlaplus.net/

6.4.2. The Zenon automatic theorem prover

Participants: Damien Doligez, David Delahaye [CNAM], Mélanie Jacquel [CNAM].

Damien Doligez continued the development of Zenon, a tableau-based prover for first-order logic with equality and theory-specific extensions. Version 0.7.1 of Zenon was released in May.

David Delahaye and Mélanie Jacquel designed and implemented (with some help from Damien Doligez) an extension of Zenon called SuperZenon, based on the Superdeduction framework of Brauner, Houtmann, and Kirchner [43].

Both Zenon and SuperZenon entered the CASC theorem-proving contest, where, as expected, SuperZenon did much better than Zenon.

6.4.3. Hybrid contract checking via symbolic simplification

Participant: Na Xu.

Program errors are hard to detect or prove absent. Allowing programmers to write formal and precise specifications, especially in the form of contracts, is one popular approach to program verification and error discovery. Na Xu formalizes and implements a hybrid contract checker for a pure subset of OCaml. The key technique we use is symbolic simplification, which makes integrating static and dynamic contract checking easy and effective. This technique statically verifies that a function satisfies its contract or blames the function violating the contract. When a contract satisfaction is undecidable, it leaves residual code for dynamic contract checking.

A paper describing this result is published in the proceeding of the PEPM'2012 conference [27]. An extended version of this paper will appear in the journal Higher-Order and Symbolic Computation. Na Xu implemented this approach in a prototype based on the OCaml 3.12.1 compiler and experimented with nontrivial examples such as sorting algorithms and balancing AVL trees (see http://gallium.inria.fr/~naxu/research/hcc.html).

6.4.4. Probabilistic contracts for component-based design

Participants: Na Xu, Gregor Goessler [project-team POPART], Alain Girault [project-team POPART].

We define a framework of probabilistic contracts for constructing component-based embedded systems, based on the formalism of discrete-time Interactive Markov Chains. A contract specifies the assumptions a component makes on its context and the guarantees it provides. Probabilistic transitions represent allowed uncertainty in the component behavior, for instance, to model internal choice or reliability. Action transitions are used to model non-deterministic behavior and communication between components. An interaction model specifies how components interact with each other.

We provide the ingredients for a component-based design flow, including (1) contract satisfaction and refinement, (2) parallel composition of contracts over disjoint, interacting components, and (3) conjunction of contracts describing different requirements over the same component. Compositional design is enabled by congruence of refinement. A paper describing the details of this result is published in the journal Formal Methods in System Design [14].

MUTANT Project-Team

6. New Results

6.1. Information-Geometric Approach to Real-time Audio Change Detection

Participants: Arnaud Dessein, Arshia Cont.

We developed a generic framework for real-time change detection of audio signals using methods of information geometry. The present method is limited to generative models of audio signals based on generic exponential distribution families. The proposed system detects changes by controlling the information rate of the signal as they arrive in time. The method also addresses shortcomings of traditional approaches based on cumulative sums which assume known parameters before change. This is achieved by calculating exact generalized likelihood ratio test statistics with complete estimation of unknown parameters in respective hypothesis [9]. The interpretation of this framework within a dually flat geometry of exponential families provide tractable algorithms for online use. Results are presented for speech segmentation into different speakers and polyphonic music segmentation.

6.2. Real-time Polyphonic Music Recognition

We investigated real-time recognition of overlapping music events in two context of dictionary-based detection and real-time alignment:

6.2.1. Real-time detection of overlapping sound events using non-negative matrix factorization

Participants: Arnaud Dessein, Arshia Cont.

Non-negative matrix factorization (NMF) methods have naturally found their way since their inception to sound and music processing. This work is an extension to our previous work in [1] on Real-time Music Transcription using sparse NMF methods. We investigate the problem of real-time detection of overlapping sound events by employing NMF techniques. We consider a setup where audio streams arrive in real-time to the system and are decomposed onto a dictionary of event templates learned off-line prior to the decomposition. An important drawback of existing approaches in this context is the lack of controls on the decomposition. We propose and compare two provably convergent algorithms that address this issue, by controlling respectively the sparsity of the decomposition and the trade-off of the decomposition between the different frequency components. Sparsity regularization is considered in the framework of convex quadratic programming, while frequency compromise is introduced by employing the beta-divergence as a cost function. The two algorithms are evaluated on the multi-source detection tasks of polyphonic music transcription, drum transcription and environmental sound recognition. The obtained results in [20] show how the proposed approaches can improve detection in such applications, while maintaining low computational costs that are suitable for real-time.

A specialized version of NMF for Real-time Music Transcription is exposed in Arnaud Dessein's PhD thesis [9].

These methods will be subject to software development in 2013.

6.2.2. Robust Real-time Polyphonic Audio-to-Score Alignment

Participant: Arshia Cont.

The *Antescofo* system is polyphonic since 2009 but its use in highly polyphonic and noisy concert environments have been challenging. To overcome this, we have studied more robust inference mechanisms. As a results, the previous inference mechanism based on maximum a posteriori of Viterbi Forward variables in mixed semi-Markov and Markov chains in [2] were abandoned in favor of a more robust method based on *importance resampling* on state-space models and smoothing of variable-order hybrid chains. This has led to robust real-time alignment and the employment of the system in various Piano performances in 2012. Further extensions are currently under study.

6.3. Real-time Multi-object Detection for Music Signals

Participants: Philippe Cuvillier [Master 2 ATIAM], Arshia Cont.

Multiple-object detection and tracking has been widely used in applications such as missile tracking and radar and has given birth to several formalisms such as Random Finite Sets [33]. Such formalisms can be seen as extensions to existing probabilistic inference mechanisms with explicit birth and death stochastic mechanisms for multiple source tracking.

In this work we aim at studying such formalisms in the case of real-time music signal processing. The idea is to track multiple sources (instruments, audio flows) from one source of observation. This approach can be beneficial to two main applications in real-time music listening:

- Extension of existing audio-to-score [2] or audio-to-audio alignment [7] mechanisms (currently based on one source) to multiple objects can address the following short-comings of existing approaches: explicit consideration for asynchrony of parallel sources; robustness to uncertainties on one or more voices.
- Studying the classical *Partial Tracking* applications in audio processing within the RFS context can lead to better results in low-level sinusoidal partial tracking of sounds.

Early studies of such formalisms are exposed in [25]. Concrete applications will be exposed in 2013.

6.4. Antescofo Language Extensions and Performance Fault-Tolerance

We have improved the *Antescofo* framework widely used for mixed instrumental and live electronic computer music. The new framework paves the way for future language extensions and paves the way for future research regarding performance fault-tolerance, synchronization mechanisms and formal verifications.

6.4.1. Antescofo Language Extensions

Participants: José Echeveste, Jean-Louis Giavitto, Florent Jacquemard, Arshia Cont.

To further extend the *Antescofo* language, the system has been formally modeled as a network of parametric timed automata in [29]. The model obtained provides operational semantics for the input scores, in particular the interaction between the instrumental and electronic parts and the timing and error handling strategies mentioned below. This approach would enable better authoring of time and interaction during programing/composing, permits to use state of the art software verification tools for the static analysis of Antescofo scores and also provides means to address critical aspects of musical performances in real-time.

In parallel, a new grammar for the score language and a new architecture have been designed for *Antescofo*, taking into account new demands from the community such as addition of timed variables in the language, dynamic time processes, time-conditional constructs, and more.

6.4.2. Performance Fault-Tolerance and Synchronization Mechanisms

Participants: José Echeveste, Jean-Louis Giavitto, Arshia Cont.

We formalized the timing strategies for musical events taking into account the variability of environment signals (musicians) and their effect on computer events programmed in *Antescofo*. The result of this work is presented in [15], where new block attributes in the language determine expected behavior in case of environment changes in real-time (errors, timing discrepancies, etc.). These additions have been implemented in the current version of the system and are widely used by the user community.

6.5. Temporal Analysis and Verification of Interactive Music Scores

Participants: Léa Fanchon [Master 2 École Centrale], Florent Jacquemard.

Léa Fanchon's Masters thesis, under the supervision of Florent Jacquemard, [26] presents an analysis module that complements the real-time score authoring and performance in *Antescofo*, with the aim of exploring possible behavior of authored programs with respect to possible deviations in human musician performance. This work employs formal methods for temporal automata networks using linear constraint inference techniques commonly in use for task scheduling and circuit verifications.

Obtained results pave the way for future works in formal verification of interactive multimedia applications, being one of the first of its kind in computer music literature, and provides the following input to programmers and artists using *Antescofo*:

- Evaluation of robustness of the program with respect to the environment's (musician's performance) temporal variations,
- Feedback to programmers/artists on critical synchronization points for better programming.

An article describing this work is currently in preparation for a submission to a computer music conference.

6.6. Formal study of Antescofo as a Reactive System

Participants: Guillaume Baudart [Master 2 ATIAM], Florent Jacquemard, Marc Pouzet [ENS], Jean-Louis Giavitto, Arshia Cont.

An *Antescofo* score/program can be considered as a specification of a reactive system through its coupling of a machine listening with a real-time synchronous language. In his master thesis under the supervision of Florent Jacquemard and Marc Pouzet (team Parkas), Guillaume Baudart has has studied the links between the reactive system of *Antescofo* and existing synchronous languages such as *Lucid Synchrone* [36] and *Reactive ML* [34]. The reactive engine of a preliminary version of *Antescofo* was developed in both languages and their structures were compared.

This study reveals the particularities of musical applications of reactive systems specific to *Antescofo* (see [24]). *Reactive ML* allows dynamic constructions but real-time performance can not be guaranteed especially when the machine listening is combined with the reactive system. On the contrary, *Lucid Synchrone* does not easily allow dynamic process creation. Each language specificity leads to strong considerations in the program/score structure for the artists. This work will be continued in 2013 to further strengthen ties between the reactive aspects of *Antescofo* and that of synchronous languages.

6.7. Tree Structured Presentation of Symbolic Temporal Data

Participants: Florent Jacquemard, Michael Rusinowitch [Project-team Cassis], Luc Segoufin [Project-team Dahu].

In traditional music notation, in particular in the languages used for the notation of mixed music such as Antescofo DSL, the durations are not expressed by numerical quantities but by symbols representing successive subdivisions of a reference time value (the beat). For this reason, trees data structures are commonly used for the symbolic representation of rhythms in computer aided composition softwares such as OpenMusic (developed at Ircam). It is therefore worth studying the applications in rhythm notation of existing formalisms for recognizing, querying, transforming and learning sets of tree structured data.

In 2012 we have studied several classes of tree recognizers which could be of interest in this context. First, with Michael Rusinowitch we have proposed in [16] a novel class of automata computing on unranked trees, which are context free in two dimensions: in the the sequence of successors of a node and also along paths. Second, we studied with Luc Segoufin [21] automata and logics computing on data trees and their relationship. Data trees are unranked ordered trees where each node carries a label from a finite alphabet and a datum from some infinite domain.

PARKAS Project-Team

6. New Results

6.1. Reactive Programming

Participants: Mehdi Dogguy, Louis Mandel, Cédric Pasteur, Marc Pouzet.

ReactiveML is an extension of OCaml with synchronous concurrency, based on synchronous parallel composition and broadcast of signals. The goal is to provide a general model of deterministic concurrency inside a general purpose functional language to program reactive systems. It is particularly suited to program discrete simulations, for instance of sensor networks.

One of the current focus of the research is being able to simulate huge systems, composed of millions of agents, by extending the current purely sequential implementation in order to be able to take advantage of multi-core and distributed architectures. This goal has led to the introduction of a new programming construct, *reactive domain*, which allows to define local time scales. These domains help for the distribution of the code but also increase the expressiveness of the language. In particular, it allows to do time refinement. A paper on this new construct and the related static analysis has been submitted. We have implemented a new runtime for ReactiveML, that uses the MPI (Message Passing Interface) library to run programs on multi-core and distributed architectures.

We have also investigated new static analyses for the language. Following the work of PhD thesis of Mehdi Dogguy, we have studied a new analysis which adds usages on signals to be able to ensure one to one communications. We have also studied a new reactivity analysis which ensures that a process can not prevent the other ones to from executing. This analysis will be published in [10].

6.2. n-Synchronous Languages

Participants: Louis Mandel [contact], Marc Pouzet, Albert Cohen, Adrien Guatto.

The n-synchronous model introduced a way to compose streams which have *almost the same clock* and can be synchronized through the use of a finite buffer.

We have designed the language Lucy-n to program in this model of computation [40]. This language is similar to the first order synchronous data-flow language Lustre in which a buffer operator is added. A dedicated type system allows to check that programs can be executed in bounded memory and to compute sufficient buffer sizes. Technically it is done through the introduction of a subtyping constraint at each bufferization point.

- In collaboration with F. Plateau (Prove&Run), we developed a new resolution constraint algorithm for the clocking of Lucy-n programs [8]. Even if the new algorithm is less efficient that the one using abstraction, it has the advantage to be more precise and thus to accept more programs. It is useful for example for the static scheduling of Latency Insensitive Designs [41].
- We worked on an extension of the synchronous model with integer clocks. This extension allows to produce and consume several values at each activation. It has large implication on the semantics, clock typing, causality and code generation of the language.
- We have continue the work on the code generation. In particular, we have been designing a new intermediate representation that allows to deal with integer clocks.

6.3. Strong normal form for large integers, boolean functions and finite automata

Participant: Jean Vuillemin.

Jean Vuillemin's recent work focusses on finding Strong Normal Form for large Integers, Boolean functions and finite Automata, with applications to circuits and software.

- [16] is the latest version of JV's course notes at ENS "De l'algorithme au circuit".
- [9] shows that the ordered dimension of a Boolean function is a lower bound on the size of most known ordered Decision Diagrams, and that ordered decision diagrams can be efficiently constructed an operated upon.
- [6] shows an approach to circuit protection against side-channel attacks based on a statistical analysis of power traces derived from actual measures of the circuit in operation.

6.4. A theory of safe optimisations in the C11/C++11 memory model and applications to compiler testing

Participants: Francesco Zappa Nardelli [contact], Robin Morisset, Pankaj Pawan.

Compilers sometimes generate correct sequential code but break the concurrency memory model of the programming language: these subtle compiler bugs are observable only when the miscompiled functions interact with concurrent contexts, making them particularly hard to detect. In this work we design a strategy to reduce the hard problem of hunting concurrency compiler bugs to differential testing of sequential code and build a tool that puts this strategy to work. Our first contribution is a theory of sound optimisations in the C11/C++11 memory model, covering most of the optimisations we have observed in real compilers and validating the claim that common compiler optisations are sound in the C11/C++11 memory model. Our second contribution is to show how, building on this theory, concurrency compiler bugs can be identified by comparing the memory trace of compiled code against a reference memory trace for the source code. Our tool identified several mistaken write introductions and other unexpected behaviours in the latest release of the gcc compiler.

A paper on this work has been submitted to an international conference [15].

6.5. A verified compiler for relaxed-memory concurrency

Participant: Francesco Zappa Nardelli [contact].

We studied the semantic design and verified compilation of a C-like programming language for concurrent shared-memory computation above x86 multiprocessors. The design of such a language is made surprisingly subtle by several factors: the relaxed-memory behaviour of the hardware, the effects of compiler optimisation on concurrent code, the need to support high-performance concurrent algorithms, and the desire for a reasonably simple programming model. In turn, this complexity makes verified (or verifying) compilation both essential and challenging. This project started in 2010, and in 2012 we submitted a journal version, describing the correctness proof of all the phases of our CompCertTSO compiler (including experimental fence eliminations). This has been accepted for publication in Journal of the ACM [3].

In collaboration with Jaroslav Sevcik (U. Cambridge), Viktor Vafeiadis (MPI-SWS), Suresh Jagannathan (Purdue U.), Peter Sewell (U. Cambridge).

6.6. Compiling C/C++ concurrency from C++11 to POWER

Participant: Francesco Zappa Nardelli [contact].

The upcoming C and C++ revised standards add concurrency to the languages, for the first time, in the form of a subtle relaxed memory model (the C++11 model). This aims to permit compiler optimisation and to accommodate the differing relaxed-memory behaviours of mainstream multiprocessors, combining simple semantics for most code with high-performance low-level atomics for concurrency libraries.

We studied the the correctness of two proposed compilation schemes for the C++11 load and store concurrency primitives to Power assembly, having noted that an earlier proposal was flawed. (The main ideas apply also to ARM, which has a similar relaxed memory architecture.)

This should inform the ongoing development of production compilers for C++11 and C1x, clarifies what properties of the machine architecture are required, and builds confidence in the C++11 and Power semantics.

A paper describing this work will appear in POPL 2012 [5].

In collaboration with Kayvan Memarian (previously student in the Moscova EPI, currently at U. Cambridge).

6.7. Compilation techniques for synchronous languages

Participants: Marc Pouzet [contact], Adrien Guatto, Léonard Gérard, Cédric Pasteur.

- The generation of efficient sequential code for synchronous data-flow languages raises two intertwined issues: control and memory optimization. While the former has been extensively studied, for instance in the compilation of Lustre and SIGNAL, the latter has been only addressed in a restricted manner. Yet, memory optimization becomes a pressing issue when arrays are added to such languages, for example, SCADE 6 [8]. We have proposed a two-levels solution to the memory optimization problem. It combines a compile-time optimization algorithm, reminiscent of register allocation, paired with language annotations on the source given by the designer. Annotations express in-place modifications and control where allocation is performed. Moreover, they allow external functions performing in-place modifications to be imported safely. Soundness of annotations is guaranteed by a semilinear type system and additional scheduling constraints. A key feature is that annotations for well-typed programs do not change the semantics of the language: removing them may lead to a less efficient code but with the very same semantics.

  The method has been implemented in HEPTAGON, the compiler developed in the team of a Lustre-like synchronous language extended with hierarchical automata and arrays. Experiments show that the proposed approach removes most of the unnecessary array copies, resulting in faster code that uses less memory. This work has been presented at the *ACM Intern. Conf. on Languages, Compilers and Tools for Embedded Systems (LCTES'12)* in June 2012 and it has received the *Best paper award*.

6.8. Generation of Parallel Code from Synchronous Programs

Participants: Albert Cohen [contact], Léonard Gérard, Adrien Guatto, Nhat Minh Le, Marc Pouzet.

- Efficiently distributing synchronous programs is a challenging and long-standing subject. This paper introduces the use of futures in a Lustre-like language, giving the programmer control over the expression of parallelism. In the synchronous model where computations are considered instantaneous, futures increase expressiveness by decoupling the beginning from the end of a computation. Through a number of examples, we show how to desynchronize long computations and implement parallel patterns such as fork-join, pipelining and data parallelism. The proposed extension preserves the main static properties of the base language, including static resource bounds and the absence of deadlock, livelock and races. Moreover, we prove that adding or removing futures preserves the underlying synchronous semantics.

  This work has been presented at the *ACM Intern. Conf. on Embedded Software (EMSOFT 2012)*, in October 2012 and it received the *Best paper award*.

  Further work along these lines is taking place, to generate code for a variety of low-overhead execution models, to cope with real-time constraints, and to formalize and prove the correctness of the underlying concurrent data structures. On the latter point, a paper has been accepted at the ACM Conf. PPoPP 2013.

6.9. Semantics and Implementation of Hybrid System Modelers

Participants: Marc Pouzet [contact], Timothy Bourke.

[8] http://www.esterel-technologies.com/products/scade-suite/

Zélus is a new programming language for modeling systems that mix discrete logical time and continuous time behaviors. From a user's perspective, its main originality is to extend an existing -like synchronous language with Ordinary Differential Equations (ODEs). The extension is conservative: any synchronous program expressed as data-flow equations and hierarchical automata can be composed arbitrarily with ODEs in the same source code. A dedicated type system and causality analysis ensure that all discrete changes are aligned with zero-crossing events so that no side effects or discontinuities occur during integration. Programs are statically scheduled and translated into sequential code which, by construction, runs in bounded time and space. Compilation is effected by source-to-source translation into a small synchronous subset which is processed by a standard synchronous compiler architecture. The resulting code is paired with an off-the-shelf numeric solver.

This experiment show that it is possible to build a modeler for explicit hybrid systems à la Simulink/Stateflow on top of an existing synchronous language, using it both as a semantic basis and as a target for code generation. In parallel with the software development done during the year, we investigate, in collaboration with Albert Benveniste, Benoit Caillaud (Inria Rennes) and Dassault-Systèmes the treatment of Differential Algebraic Equations (DAEs), in explicit or semi-explicit form.

This work will be presented at the *ACM Intern. Conference on Hybrid Systems: Computation and Control (HSCC 2013)* in April 2013.

PI.R2 Project-Team

6. New Results

6.1. Proof-theoretical and effectful investigations

Participants: Federico Aschieri, Pierre Boutillier, Pierre-Louis Curien, Hugo Herbelin, Guillaume Munch-Maccagnoni, Pierre-Marie Pédrot, Alexis Saurin, Arnaud Spiwack.

6.1.1. Sequent calculus and computational duality

System L syntax. Pierre-Louis Curien studied in some detail the differences (and translations) between variants of "system L" syntax for polarised classical logic (developed by Guillaume Munch-Maccagnoni and himself):

- weakly focalised systems (where negatives can be worked on at any moment in a proof) versus focalised systems (where negative and positive phases alternate strictly), versus strongly focalised systems (where furthermore negative phases have to decompose negatives completely);
- systems where changes of polarity are implicit (like in Girard's LC) versus systems where they are explicitly marked using shift operators. These shift operators are formally adjoint, and as a matter of fact a suitable intuitionistic fragment of system L corresponds exactly to Levi's CBPV;
- systems with stoup (which retain only proofs that follow the focalisation discipline) versus (still focalised) systems without stoup (where the focalisation is forced by the dynamics of reduction);
- one-sided systems (with an implicit negation given by De Morgan duality) versus two-sided systems (allowing for explicit negation, and for distinguishing the left/right and the positive/negative dualities).

Pierre-Louis Curien is also currently studying a polarised version of a notion of general connective suggested earlier by Hugo Herbelin (unpublished work), and the composition structure of these connectives (in the spirit of operads).

Categorical semantics. Guillaume Munch-Maccagnoni investigated a notion of "direct style" for adjunction models, inspired by his work on polarisation in the "L" system, in the lineage of Führmann's [47] direct-style characterisation of monadic models. (It is part of joint work with Marcelo Fiore and Pierre-Louis Curien.)

Polarised Peano arithmetic. Guillaume Munch-Maccagnoni investigates the computational contents of polarised classical logic in arithmetic and in natural deduction. This allows him to compare the constructivisation of the principle $\neg\forall \Rightarrow \exists\neg$ based on classical realisability (Krivine) and the one based on delimited control (via "double negation shift"); both of which seem to be simplified by a better understanding of the "formulae-as-types" paradigm for a negation which is involutive in a strong sense.

Guillaume Munch-Maccagnoni investigates how a notion of classical realisability structure (inspired by Krivine's) can be used to prove properties of type systems which are usually regarded as syntactic.

Classical call-by-need and the duality of computation. In 2011, Zena Ariola, Hugo Herbelin and Alexis Saurin characterised the semantics of call-by-need calculus with control in the framework of *the duality of computation*. The same set of authors extended with Paul Downen and Keiko Nakata worked on abstract machines and continuation-passing-style semantics for call-by-need with control, resulting into a paper presented at FLOPS 2012 [20].

Further work has been done by Zena Ariola, Hugo Herbelin, Luís Pinto, Keiko Nakata and José Espírito Santo on typing the continuation-passing-style of call-by-need calculus, opening the way to a proof of normalisation of simply-typed call-by-need with control, and from there to a proof of consistency of classical arithmetic with dependent choice.

Zena Ariola also investigated how to formulate a parametric theory which encompasses call-by-value, call-by-name and call-by-need. Each theory is obtained by giving the appropriate definition of what is a value and a co-value. The theory also includes so called lifting axioms which allow one to relax the syntactic restrictions previously imposed on the call-by-value, call-by-name and call-by-need calculi. The theory also allows to include the η-rules which before were causing confluence to fail. The approach can be applied to natural deduction and this allows to express different embeddings of natural deduction into sequent calculus directly in the theory. The advantage of the new formalisation is that analogously to natural deduction, one can experiment with different strategies starting from the same term. Moreover, the theory is well-suited for continuation passing style transformation and, in particular, it leads to a different and simpler formalisation of classical call-by-need, its abstract machine and continuation passing style.

6.1.2. Dependent monads

Pierre-Marie Pédrot generalised the notion of monad in order to be able to use it in a dependent framework. This new structure allows to write effects in a pure functional language, such as Coq, through a monadic encoding.

This way, the whole monadic apparatus can be lifted to dependent programs, as well as proofs.

6.1.3. Linear dependent types

Arnaud Spiwack continued his investigations on dependently typed linear sequent calculus (based on Curien & Herbelin's $\mu\tilde{\mu}$). The current version of his system resembles Andreoli's focalised linear logic (yet to be published).

Pierre-Marie Pédrot has been working on a delimited CPS translation of the Calculus of Inductive Constructions, seen through the prism of polarised linear logic. Restricting dependencies to positives naturally fits into the scenery of delimited control, while extending negatives to infinitary objects permits to recover some properties of the involutivity of linear double-negation.

6.1.4. Proving with side-effects

Axiom of dependent choice. 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. This work has been presented at LICS 2012 [22].

Memory assignment, forcing and delimited control. Hugo Herbelin investigated how to extend his work on intuitionistically proving Markov's principle [54] and the work of Danko Ilik on intuitionistically proving the double negation shift (i.e. $\forall x\,\neg\neg A \to \neg\neg\forall x\,A$) [15] to other kind of effects. In particular, memory assignment is related to Cohen's forcing as emphasised by Krivine [58] and by the observation that Cohen's translation of formula P into $\forall y \leq x \exists z \leq y\, P(z)$ is similar to a state-passing-style transformation of type P into $S \to S \times P$.

Hugo Herbelin then designed a logical formalism with memory assignment that allows to *prove* in direct-style any statement provable using the forcing method, the same way as logic extended with control operators allows to support direct-style classical reasoning. Thanks to the use of delimiters over "small" formulas similar to the notation of Σ^0_1-formulas in arithmetic, the whole framework remains intuitionistic, in the sense that it satisfies the disjunction and existence property.

Two typical applications of proving with side-effects are global-memory proofs of the axiom of countable choice and an enumeration-free proof of Gödel's completeness theorem.

The main ideas of this research program have been communicated during the Logic and Interaction weeks in Marseille in February 2012.

In the continuation of his work with Silvia Ghilezan [4] on showing that Saurin's variant $\Lambda\mu$ [8] of Parigot's $\lambda\mu$-calculus [65] for classical logic was a canonical call-by-name version of Danvy-Filinski's call-by-value calculus of delimited control, Hugo Herbelin studied with Alexis Saurin and Silvia Ghilezan another variant of call-by-name calculus of delimited control. This is leading to a general paper on call-by-name and call-by-value delimited control.

Classical logic, stack calculus and stream calculus. Alexis Saurin studied the connection between the stack calculus recently proposed by Ehrhard et al and $\lambda\mu$-calculus and how the former can be precisely compared to the target of the CPS of the latter. He also investigated separation issues related to the stack calculus. During a visit to UPenn in the spring, Alexis Saurin and Marco Gaboardi investigated type systems for a stream calculus which contains $\Lambda\mu$.

Moreover, Alexis Saurin's paper *Böhm theorem and Böhm trees for the $\Lambda\mu$-calculus* [16] was published in TCS early 2012.

6.1.5. PTS and delimited control

From the study of one-pass CPS on the one side and of previous presentations of pure type systems with control operators on the other side, Pierre Boutillier and Hugo Herbelin have investigated how splitting terms into categories opens a new way to merge dependent types and monads. A preliminary set of rules has been presented during the third week of Logic and Interaction in Marseille.

It was refined since then but has not reached yet the maturity required to be accepted for publication in an international conference.

6.1.6. Interactive realisability

Thanks to the Curry-Howard correspondence for classical logic, it is possible to extract programs from classical proofs. These programs use control operators as a way to implement backtracking and processes of intelligent learning by trial and error. Unfortunately, such programs tend to be poorly efficient. The reason is that, in a sense, they are designed in order to keep their correctness and termination proofs simple. Each small modification of these programs seems, at best, to require major and difficult adaptations of their correctness proofs. This is due to a lack of understanding and control of the backtracking mechanism that interprets classical proofs. In order to write down more efficient programs, it is necessary to describe exactly: a) what the programs learn, b) how the knowledge of programs varies during the execution.

A first step towards this goal is the theory of Interactive Realisability, a semantics for intuitionistic arithmetic with excluded middle over semi-decidable predicates. It is based on a notion of state, which describes the knowledge of programs coming from a classical proof, and explains how the knowledge evolves during computation.

Federico Aschieri has extended the theory of interactive realisability to a full classical system, namely first-order Peano arithmetic with Skolem axioms. This is a very expressive system, with non-trivial axioms of choice and comprehension. The resulting programs are interpreted as stratified-learning algorithms, which build in a very organised way the approximations of the Skolem functions used in the proofs. The work has appeared in the proceedings of the conference Computer Science Logic 2012. A careful implementation of this extended theory –yet to be developed – will lead to a dramatic efficiency improvement over the already existing computational interpretations.

Federico Aschieri has also showed how to use interactive realisability to provide purely proof-theoretic results. He proved with a new method the conservativity of Peano arithmetic with Skolem axioms over Peano arithmetic alone for arithmetical formulas. In particular, the method can be seen as a constructivisation and substantial refinement of Avigad's forcing. The work has appeared in the proceedings of the workshop Classical Logic and Computation 2012.

6.1.7. Reverse mathematics

Hugo Herbelin explored with Gyesik Lee and Keiko Nakata the constructive content of the big five subsystems of second-order arithmetic considered in the context of (classical) reverse mathematics. They obtained a

uniform characterisation of these systems in terms of variants of the comprehension axiom called separation, co-separation and interpolation.

This is the first step in a larger project attempting first to connect to predicative type theory the subsystems of System F underlying the proof-as-program structure of the big five subsystems of second-order arithmetic, and secondly to reformulate these subsystems in terms of pure systems of inductive definitions.

Jaime Gaspar has several projects running simultaneously. For example, in one of his projects he created a small unoptimised automated theorem prover, and he hopes to optimise it and use it to obtain a certain completely formalised proof to which he can apply a proof interpretation in order to extract computational content. As another example, in another project he his trying to show that several classical proof interpretations are instances of a unified proof interpretation, in a parallel way to what is known for intuitionistic proof interpretations.

6.2. Type theory and the foundations of Coq

Participants: Pierre Boutillier, Pierre-Louis Curien, Hugo Herbelin, Pierre-Marie Pédrot, Yann Régis-Gianas, Alexis Saurin, Matthieu Sozeau.

6.2.1. Calculus of inductive constructions and typed equality

The work of Hugo Herbelin and Vincent Siles on the equivalence of Pure Type Systems with typed or untyped equality has been published [17].

6.2.2. Substitutions and isomorphisms

Pierre-Louis Curien completed his joint work with Martin Hofmann (Univ. of Munich) and Richard Garner (MacQuarie University, Sydney) on comparing two categorical interpretations of (extensional) type theories. More precisely, we wanted to compare two ways of giving a categorical interpretation of Martin-Löf type theory, both overcoming the following mismatch: syntax has exact substitutions, while their categorical interpretation, in terms of pullbacks or fibrations, "implements" substitutions only up to isomorphism. One can then either change the model (strictification) [55], or modify the syntax (by introducing explicit substitutions and more importantly explicit coercions between types that are now only isomorphic) [2]. In the latter case, one has to prove a coherence theorem to show that the interpretation is in the end independent from these coercion decorations. Such a proof was given in [2], using rewriting methods. These approaches turn out to be related through a general machinery that relates three kinds of categories, with strict or non strict objects and morphisms. As a bonus we get a new, more conceptual proof of coherence. These results are now being written up for a special issue in honour of Glynn Winskel. In further work, we wish to address intensional, and homotopy type-theoretic versions of these coherence problems.

6.2.3. Homotopy type theory

The univalence axiom proposed by Voevodsky states that for any two types to be equal exactly means being of same cardinality. This new axiom for type theory turns to have very interesting consequences for the practical foundations of formal mathematical reasoning: it smoothly implies other axioms such as functional extensionality or propositional existential but before all it says that any property proved about some mathematical structure immediately applies to any other other type ("sets" informally) which it is isomorphic to.

This axiom however contradicts the current logical foundations of Coq (in the presence of Streicher's axiom K). Investigations have then been started to understand how to weaken the Calculus of Inductive Constructions implemented in Coq so as to make it compatible with univalence. In a first step, this resulted in the design of a new rule for singleton elimination that has been implemented by Hugo Herbelin as an optional feature of Coq (singleton elimination is the ability to build objects in datatypes from canonically-proved propositional properties such as equality).

6.2.4. Models of type theory

The existing models of homotopy type theory are based on simplicial sets or on their extensions as Kan complexes. 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. The technique he used seems to straightforwardly generalise to provide type-theoretic constructions for arbitrary presheaves on inductively generated categories.

6.2.5. Forcing in type theory

Together with Nicolas Tabareau and Guilhem Jaber (Inria Ascola team, École des mines de Nantes), Matthieu Sozeau investigated an internalisation of the presheaf model of the Calculus of Inductive Constructions (CIC). They published their work at LICS'12 [23]. This work corresponds to adapting the idea of Forcing due to Cohen in Type Theory. An internal model construction allows to enrich the logical type theory with new modalities and define their semantics by translation to CIC. The usual Cohen forcing can be realised using this framework to show the independence of the continuum hypothesis in CIC, but more practical applications are possible as well. Notably, the step-indexed technique for building models of imperative languages with rich type structure can be phrased as a forcing/presheaf construction. Sozeau, Tabareau and Jaber developed a plugin that can handle this example [32] which relies on a modified Coq version implementing proof-irrelevance and eta-rules for records.

6.2.6. Proof irrelevance, eta-rules

Matthieu Sozeau continued his work on proof-irrelevance by implementing a variant of Werner's proof-irrelevant CIC in Coq [72]. An article describing this work is in preparation. The new system also handles the extensional eta-rules for records, extending the technique implemented by Hugo Herbelin to handle eta-expansion of functions in Coq.

6.2.7. Unification

The unification algorithm of Coq now essentially dwells in the λ-calculus part of the language. Pierre Boutillier started a refactoring of the code in order to deal with algebraic datatypes. Hugo Herbelin and Pierre Boutillier investigated how to reformulate unification on top of an abstract machine (i.e. on top of sequent calculus). Hugo Herbelin added various heuristics to the unification algorithm of Coq, making them both more powerful and customisable.

Matthieu Sozeau is continuing work in collaboration with Beta Ziliani (PhD student of Derek Dreyer at MPI Saarbrücken, two one week visits in 2012), and Aleksandar Nanevski (Researcher at IMDEA Madrid) on giving a clear formalisation for the unification algorithm of Coq. This will help understand better the working of advanced features like Canonical Structures and Type Classes that are heavily used in big developments, as the spectacular recently completed formalisation of the proof of Feit-Thompson's Odd theorem by the Mathematical Components team.

Matthieu Sozeau adapted the existing unification algorithm to be universe-aware, resulting in more predictability and earlier error-reporting in both the type inference and tactic unification algorithms of Coq.

6.3. Homotopy of rewriting systems

Participants: Pierre-Louis Curien, Yves Guiraud, Philippe Malbos.

6.3.1. Coherence in monoidal structures

Yves Guiraud and Philippe Malbos have applied the Squier's homotopical theorem [70], which they had generalised to higher-dimensional rewriting systems [52], to several types of categories with monoidal structures. This work develops a formal setting to produce constructive proofs of coherence conditions, applied to the cases of monoidal categories, symmetric monoidal categories and braided categories. These results have been published in Mathematical Structures in Computer Science [12].

6.3.2. Computation of resolutions of monoids

Yves Guiraud and Philippe Malbos have extended Squier's homotopical theory to the higher dimensions of presentations of monoids to get an algorithm transforming a convergent word rewriting system into a polygraphic resolution of the presented monoid, in the setting of Métayer [63]. From this polygraphic resolution, this work gives an explicit procedure to recover several of the known Abelian resolutions of the monoid, generalising and relating algebraic invariants of monoids. Moreover, a polygraphic resolution corresponds to the normalisation strategies of rewriting systems and they contain all the proofs of equality between elements, together with the proofs of equality of those proofs of equality, and so on. This work has been published in Advances in Mathematics [13]. By nature, polygraphic resolutions bear many similarities with the higher-dimensional groupoids that appear in homotopical type theory when one considers the towers of identity types: this connection will be investigated by Pierre-Louis Curien, Yves Guiraud, Hugo Herbelin and Matthieu Sozeau.

6.3.3. Coherent presentations of Artin groups

With Stéphane Gaussent (Institut Camille Jordan, Université de Saint-Étienne), Yves Guiraud and Philippe Malbos are currently finishing an article on the rewriting properties and coherence issues in Artin groups, a class of groups 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 group [71], and one by Deligne about the actions of Artin groups on categories [44], both proved by geometrical and non-constructive 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 of both those theorems. In fact, the method even improves the formerly known results: for example, it generalises Deligne's result to cases where his geometrical proof cannot be applied. A preliminary version of this work is available online [31]. The next objective of this collaboration is to extend the formal setting and methods to compute polygraphic resolutions of Artin groups, with a view towards two open problems of combinatorial group theory with respect to Artin groups: the decidability of the word problem and the verification that a precise topological space is a classifying space.

6.3.4. Higher-dimensional linear rewriting

With Samuel Mimram (CEA Saclay) and Pierre-Louis Curien, Yves Guiraud and Philippe Malbos investigate the links between set-theoretic rewriting theory and the computational methods known in symbolic algebra, such as Gröbner bases [36]. In particular, this work is interested in extending the setting of higher-dimensional rewriting to include "linear rewriting" and, as a consequence, to be able to apply its methods in symbolic computation. One particular direction is to understand Anick's resolution [33], and to improve it with the completion-reduction methodology, in order to get better algorithms to compute homological invariants and to prove important properties such as Koszulness. This research direction has been presented to the first call for projects of the IDEX Sorbonne-Paris-Cité, together with Eric Hoffbeck and Muriel Livernet (LAGA, Université Paris 13) and François Métayer (PPS, Université Paris 7).

6.4. Coq as a functional programming language

Participants: Nicolas Ayache, Pierre Boutillier, François Bobot, Guillaume Claret, Lourdes del Carmen Gonzalez Huesca, Tim Griffin, Hugo Herbelin, Pierre Letouzey, Matthias Puech, Yann Régis-Gianas, Matthieu Sozeau.

6.4.1. Type classes and libraries

Pierre Castéran from Inria Bordeaux and Matthieu Sozeau published a tutorial on the use of type classes [30] that was used as the basis of an invited lecture by M. Sozeau at the JFLA conference in February 2012. It will be published as part of the new version of the Coq'Art book.

6.4.2. Dependent pattern-matching

Pierre Boutillier experimented about how to integrate gently in Coq the compilation process he came up with last year to simulate Agda-style dependent pattern-matching. As a consequence, pattern grammar in Gallina has changed, much more notations can be used and users can write patterns instead of simple abstractions in the pattern-matching return clause.

Matthieu Sozeau continued maintenance and polishing of the Equations plugin that allows dependent pattern-matching on inductive families. A first official release is planned for the beginning of 2013.

6.4.3. Modularised arithmetical libraries

The modularised arithmetical libraries elaborated by Pierre Letouzey during the previous year(s) have been released as part of Coq 8.4. They provided greater uniformity of available functions and lemmas across the various Coq numerical datatypes. These libraries seem to work quite well, the only remaining issue is the documentation: due to this complex modular organisation, it is currently tedious for the user to browse the available functions and results. We expect to tackle this issue next year, by providing various documentation views, either the external summary of all available elements at once, or the internal layout of these elements.

6.4.4. Library of finite sets

Pierre Letouzey has integrated an additional Coq implementation in the MSets library of finite sets. This additional implementation is an improved version of the Red-Black-Tree library contributed by Andrew Appel. Using these RBT instead of the previously available AVL could be more efficient, at least in Coq, since they trigger no computations of integer numbers coding the tree depth.

6.4.5. Library on XPath processing

As part of the ANR Typex (http://typex.lri.fr), Matthieu Sozeau is developing a library for the certification of efficient XPath/XQuery engines in collaboration with Kim Nguyen (LRI) and Alan Schmitt (Inria Grenoble).

6.4.6. Mathematics of routing

Tim Griffin's primary focus during his visit to πr^2 was the development of a "metalanguage" for algebraic structures using Coq. Since he was something of a beginner with Coq, this involved learning the basics as well as more advanced work on representing algebraic structures. He made very good progress on this while in Paris and is now continuing this work in Cambridge.

6.4.7. Incrementality in proof languages

Matthias Puech and Yann Régis-Gianas worked on incremental type checking. This preliminary work has been presented during a contributed talk at TLDI 2012 [25]. It sets the ground for an incremental proof development and checking system, by means of a representation language for repositories of proofs and proof changes.

The traditional interaction with a proof-checker is a batch process. Coq (among others) refines this model by providing a read-eval print loop with a global undo system, implemented in an ad-hoc way. A more general approach to incrementality is being developed by means of a finer-grained analysis of dependencies. The approach developed is adaptable to any typed formal language: the language is specified in a meta-language close to the Edinburgh Logical Framework, in which subsequent versions of a development can be type-checked incrementally. Applications of this framework are: proof language for proof assistants, integrated development environments for proof or programming languages, typed version control systems.

6.4.8. Proofs of programs in Coq

As part of the CerCo European project, in collaboration with Roberto Amadio (PPS, Paris 7), François Bobot, Nicolas Ayache and Yann Régis-Gianas maintained a prototype compiler for a large subset of the C language whose specificity is to annotate input source programs with information about the worst-case execution cost of their machine code. Yann Régis-Gianas started a mechanised version of the proof technique used to prove the correctness of such an annotating compiler.

Yann Régis-Gianas maintained another compiler for Core ML that uses a generalisation of CerCo technique in order to obtain certified worst case execution time bounds on functional programs. This compiler produces proof obligations in Coq. The corresponding paper is published in January 2012 in the proceedings of FOPARA 2011 [19].

Nicolas Ayache developed a Frama-C plugin distributed in the CerCo software suite whose role is to synthesize cost annotations out of C programs. François Bobot developed a new version of this plugin. In particular, this new version handles C programs that manipulate pointers.

In collaboration with Roberto Amadio, Yann Régis-Gianas extended the cost annotating compilation chain of the FOPARA paper to handle the cost of memory management. A journal paper is about to be published in HOSC.

6.4.9. Lightweight proof-by-reflection

In the context of the ANR project Paral-ITP, Lourdes del Carmen Gonzalez Huesca, Guillaume Claret and Yann Régis-Gianas developed a new technique for proof-by-reflection based on a notion of *a posteriori* simulation of effectful computations in Coq.

POLSYS Project-Team

6. New Results

6.1. The complexity of solving quadratic boolean systems is better than exhaustive search

A fundamental problem in computer science is to find all the common zeroes of m quadratic polynomials in n unknowns over $\mathbb{F}_2$. The cryptanalysis of several modern ciphers reduces to this problem. Up to now, the best complexity bound was reached by an exhaustive search in $4\log_2 n\, 2^n$ operations. In [4], we give an algorithm that reduces the problem to a combination of exhaustive search and sparse linear algebra. This algorithm has several variants depending on the method used for the linear algebra step. Under precise algebraic assumptions on the input system, we show in [4] that the deterministic variant of our algorithm has complexity bounded by $O(2^{0.841n})$ when $m = n$, while a probabilistic variant of the Las Vegas type has expected complexity $O(2^{0.792n})$. Experiments on random systems show that the algebraic assumptions are satisfied with probability very close to 1. We also give a rough estimate for the actual threshold between our method and exhaustive search, which is as low as 200, and thus very relevant for cryptographic applications.

6.2. Improving the Complexity of Index Calculus Algorithms in Elliptic Curves over Binary Fields

In [25], we study the index calculus method that was first introduced by Semaev for solving the ECDLP and later developed by Gaudry and Diem. In particular, we focus on the step which consists in decomposing points of the curve with respect to an appropriately chosen factor basis. This part can be nicely reformulated as a purely algebraic problem consisting in finding solutions to a multivariate polynomial. Our main contribution is the identification of particular structures inherent to such polynomial systems and a dedicated method for tackling this problem. We solve it by means of Gröbner basis techniques and analyze its complexity using the multi-homogeneous structure of the equations. We emphasize that the complexity obtained in the paper is very conservative in comparison to experimental results. We hope the new ideas provided here may lead to efficient index calculus based methods for solving ECDLP in theory and practice.

6.3. On the relation between the MXL family of algorithms and Gröbner basis algorithms

The computation of Gröbner bases remains one of the most powerful methods for tackling the Polynomial System Solving (PoSSo) problem. The most efficient known algorithms reduce the Gröbner basis computation to Gaussian eliminations on several matrices. However, several degrees of freedom are available to generate these matrices. It is well known that the particular strategies used can drastically affect the efficiency of the computations. In this work, we investigate a recently-proposed strategy, the so-called "Mutant strategy", on which a new family of algorithms is based (MXL, MXL2 and MXL3). By studying and describing the algorithms based on Gröbner basis concepts, we demonstrate in [3] that the Mutant strategy can be understood to be equivalent to the classical Normal Selection Strategy currently used in Gröbner basis algorithms. Furthermore, we show that the "partial enlargement" technique can be understood as a strategy for restricting the number of S-polynomials considered in an iteration of the F4 Gröbner basis algorithm, while the new termination criterion used in MXL3 does not lead to termination at a lower degree than the classical Gebauer–Möller installation of Buchberger's criteria. We claim that our results map all novel concepts from the MXL family of algorithms to their well-known Gröbner basis equivalents. Using previous results that had shown the relation between the original XL algorithm and F4, we conclude that the MXL family of algorithms can be fundamentally reduced to redundant variants of F4.

6.4. On the Complexity of the BKW Algorithm on LWE

In [35], 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, as a result, provide new upper bounds for the concrete hardness of these LWE-based schemes.

6.5. On the Complexity of the Arora-Ge algorithm against LWE

Arora & Ge recently showed that solving LWE can be reduced to solve a high-degree non-linear system of equations. They used a linearization to solve the systems. We investigate in [34] the possibility of using Gröbner bases to improve Arora & Ge approach.

6.6. On enumeration of polynomial equivalence classes and their application to MPKC

The Isomorphism of Polynomials (IP) is one of the most fundamental problems in multivariate public key cryptography (MPKC). In [8], we introduce a new framework to study the counting problem associated to IP. Namely, we present tools of finite geometry allowing to investigate the counting problem associated to IP. Precisely, we focus on enumerating or estimating the number of isomorphism equivalence classes of homogeneous quadratic polynomial systems. These problems are equivalent to finding the scale of the key space of a multivariate cryptosystem and the total number of different multivariate cryptographic schemes respectively, which might impact the security and the potential capability of MPKC. We also consider their applications in the analysis of a specific multivariate public key cryptosystem. Our results not only answer how many cryptographic schemes can be derived from monomials and how big the key space is for a fixed scheme, but also show that quite many HFE cryptosystems are equivalent to a Matsumoto-Imai scheme.

6.7. Cryptanalysis of HFE, Multi-HFE and Variants for Odd and Even Characteristic

In [5], we investigate the security of HFE and Multi-HFE schemes as well as their minus and embedding variants. Multi-HFE is a generalization of the well-known HFE schemes. The idea is to use a multivariate quadratic system - instead of a univariate polynomial in HFE - over an extension field as a private key. According to the authors, this should make the classical direct algebraic (message-recovery) attack proposed by Faugère and Joux on HFE no longer efficient against Multi-HFE. We consider here the hardness of the key-recovery in Multi-HFE and its variants, but also in HFE (both for odd and even characteristic). We first improve and generalize the basic key recovery proposed by Kipnis and Shamir on HFE. To do so, we express this attack as matrix/vector operations. In one hand, this permits to improve the basic Kipnis-Shamir (KS) attack on HFE. On the other hand, this allows to generalize the attack on Multi-HFE. Due to its structure, we prove that a Multi-HFE scheme has much more equivalent keys than a basic HFE. This induces a structural weakness which can be exploited to adapt the KS attack against classical modifiers of multivariate schemes such as minus and embedding. Along the way, we discovered that the KS attack as initially described cannot be applied against HFE in characteristic 2. We have then strongly revised KS in characteristic 2 to make it work. In all cases, the cost of our attacks is related to the complexity of solving MinRank. Thanks to recent complexity results on this problem, we prove that our attack is polynomial in the degree of the extension field for all possible practical settings used in HFE and Multi-HFE. This makes then Multi-HFE less secure than basic HFE for equally-sized keys. As a proof of concept, we have been able to practically break the most conservative proposed parameters of multi-HFE in few days (256 bits security broken in 9 days).

6.8. Solving Polynomial Systems over Finite Fields: Improved Analysis of the Hybrid Approach

The Polynomial System Solving (PoSSo) problem is a fundamental NP-Hard problem in computer algebra. Among others, PoSSo have applications in area such as coding theory and cryptology. Typically, the security of multivariate public-key schemes (MPKC) such as the UOV cryptosystem of Kipnis, Shamir and Patarin is directly related to the hardness of PoSSo over finite fields. The goal of [22] is to further understand the influence of finite fields on the hardness of PoSSo. To this end, we consider the so-called *hybrid approach*. This is a polynomial system solving method dedicated to finite fields proposed by Bettale, Faugère and Perret (Journal of Mathematical Cryptography, 2009). The idea is to combine exhaustive search with Gröbner bases. The efficiency of the hybrid approach is related to the choice of a trade-off between the two methods. We propose here an improved complexity analysis dedicated to quadratic systems. Whilst the principle of the hybrid approach is simple, its careful analysis leads to rather surprising and somehow unexpected results. We prove that the optimal trade-off (i.e. number of variables to be fixed) allowing to minimize the complexity is achieved by fixing a number of variables proportional to the number of variables of the system considered, denoted n. Under some natural algebraic assumption, we show that the asymptotic complexity of the hybrid approach is $2^{(3.31-3.62\ \log_2(q)^{-1})\,n}$, where q is the size of the field (under the condition in particular that $\log(q) \ll n$). This is to date, the best complexity for solving PoSSo over finite fields (when $q > 2$). We have been able to quantify the gain provided by the hybrid approach compared to a direct Gröbner basis method. For quadratic systems, we show (assuming a natural algebraic assumption) that this gain is exponential in the number of variables. Asymptotically, the gain is $2^{1.49\,n}$ when both n and q grow to infinity and $\log(q)$.

6.9. Efficient Arithmetic in Successive Algebraic Extension Fields Using Symmetries

In [15] we present new results for efficient arithmetic operations in a number field $\mathbb{K}$ represented by successive extensions. These results are based on multi-modular and evaluation–interpolation techniques. We show how to use intrinsic symmetries in order to increase the efficiency of these techniques. Applications to splitting fields of univariate polynomials are presented.

6.10. Algebraic Crypanalysis with Side Channels Information

In [6] and [24] (see also the PhD thesis of C. Goyet [1]), we present new cryptanalyses of symmetric and asymmetric cryptosystems (e.g. AES and ECDSA). These analyses share the same fundamental hypotheses that some information are provided to the attacker by some oracle. In a practical point of view, such an oracle can be represented as a partial side channel attack realized in a first step (e.g. SPA, Fault attacks). The second step of the attack uses algorithms from computer algebra (e.g. Gröbner basis computation, LLL) in order to retrieve the secret key.

6.11. Worst case complexity of the Continued Fraction (CF) algorithm.

In [16] we consider the problem of isolating the real roots of a square-free polynomial with integer coefficients using the classic variant of the continued fraction algorithm (CF), introduced by Akritas. We compute a lower bound on the positive real roots of univariate polynomials using exponential search. This allows us to derive a worst case bound of $\widetilde{O}(d^4\tau^2)$ for isolating the real roots of a polynomial with integer coefficients using the *classic variant of CF*, where d is the degree of the polynomial and τ the maximum bitsize of its coefficients. This improves the previous bound of Sharma by a factor of d^3 and matches the bound derived by Mehlhorn and Ray for another variant of CF which is combined with subdivision; it also matches the worst case bound of the classical subdivision-based solvers STURM, DESCARTES, and BERNSTEIN.

6.12. Local Generic Position for Root Isolation of Zero-dimensional Triangular Polynomial Systems.

In [30] we present an algorithm based on local generic position (LGP) to isolate the complex or real roots and their multiplicities of a zero-dimensional triangular polynomial system. The Boolean complexity of the algorithm for computing the real roots is single exponential: $\widetilde{O}_B(N^{n^2})$, where $N = \max\{d, \tau\}$, d and τ, is the degree and the maximum coefficient bitsize of the polynomials, respectively, and n is the number of variables.

6.13. Univariate Real Root Isolation in Multiple Extension Fields

In [31] we present algorithmic, complexity and implementation results for the problem of isolating the real roots of a univariate polynomial in $K[x] \in L[y]$, where $L = \mathbb{Q}(\alpha\_1, \cdots, \alpha\_\ell)$ is an algebraic extension of the rational numbers. Our bounds are single exponential in ℓ and match the ones presented for the case $\ell = 1$. We consider two approaches. The first, indirect approach, using multivariate resultants, computes a univariate polynomial with integer coefficients, among the real roots of which are the real roots of B_α. The Boolean complexity of this approach is $\widetilde{O}_B(N^{4\ell+4})$, where N is the maximum of the degrees and the coefficient bitsize of the involved polynomials. The second, direct approach, tries to solve the polynomial directly, without reducing the problem to a univariate one. We present an algorithm that generalizes Sturm algorithm from the univariate case, and modified versions of well known solvers that are either numerical or based on Descartes' rule of sign. We achieve a Boolean complexity of $\widetilde{O}_B(\min\{N^{4\ell+7}, N^{2\ell^2+6}\})$ and $\widetilde{O}_B(N^{2\ell+4})$, respectively. We implemented the algorithms in C as part of the core library of Mathematica and we illustrate their efficiency over various data sets.

6.14. Mixed volume and distance geometry techniques for counting Euclidean embeddings of rigid graphs.

A graph G is called generically minimally rigid in $\mathbb{R}^d$ if, for any choice of sufficiently generic edge lengths, it can be embedded in $\mathbb{R}^d$ in a finite number of distinct ways, modulo rigid transformations. In [37] we deal with the problem of determining tight bounds on the number of such embeddings, as a function of the number of vertices. The study of rigid graphs is motivated by numerous applications, mostly in robotics, bioinformatics, and architecture. We capture embeddability by polynomial systems with suitable structure, so that their mixed volume, which bounds the number of common roots, yields interesting upper bounds on the number of embeddings. We explore different polynomial formulations so as to reduce the corresponding mixed volume, namely by introducing new variables that remove certain spurious roots, and by applying the theory of distance geometry. We focus on $\mathbb{R}^2$ and $\mathbb{R}^3$, where Laman graphs and 1-skeleta of convex simplicial polyhedra, respectively, admit inductive Henneberg constructions. Our implementation yields upper bounds for $n \leq 10$ in $\mathbb{R}^2$ and $\mathbb{R}^3$, which reduce the existing gaps and lead to tight bounds for $n \leq 7$ in both $\mathbb{R}^2$ and $\mathbb{R}^3$; in particular, we describe the recent settlement of the case of Laman graphs with 7 vertices. We also establish the first lower bound in $\mathbb{R}^3$ of about 2.52^n, where n denotes the number of vertices.

6.15. Variant Quantifier Elimination

In [10], we describe an algorithm (VQE) for a *variant* of the real quantifier elimination problem (QE). The variant problem requires the input to satisfy a certain *extra condition*, and allows the output to be *almost* equivalent to the input. The motivation/rationale for studying such a variant QE problem is that many quantified formulas arising in applications do satisfy the extra conditions. Furthermore, in most applications, it is sufficient that the output formula is almost equivalent to the input formula. The main idea underlying the algorithm is to substitute the repeated projection step of CAD by a single projection without carrying out a parametric existential decision over the reals. We find that the algorithm can tackle important and challenging problems, such as numerical stability analysis of the widely-used MacCormack's scheme. The problem has been practically out of reach for standard QE algorithms in spite of many attempts to tackle it. However the current implementation of VQE can solve it in about 12 hours.

6.16. Global optimization

Let $f_1, \cdots, f_p$ be in $\mathbb{Q}[\mathbf{X}]$, where $\mathbf{X} = (X_1, \cdots, X_n)^t$, that generate a radical ideal and let V be their complex zero-set. Assume that V is smooth and equidimensional. Given $f \in \mathbb{Q}[X]$ bounded below, consider the optimization problem of computing $f^\star = \inf_{x \in V \cap \mathbb{R}^n} f(x)$. For $\mathbf{A} \in GL_n(\mathbb{C})$, we denote by $f^{\mathbf{A}}$ the polynomial $f(\mathbf{AX})$ and by $V^{\mathbf{A}}$ the complex zero-set of $f_1^{\mathbf{A}}, ..., f_p^{\mathbf{A}}$. In [9], we construct families of polynomials $\mathsf{M}_0^{\mathbf{A}}, ..., \mathsf{M}_d^{\mathbf{A}}$ in $\mathbb{Q}[\mathbf{X}]$: each $\mathsf{M}_i^{\mathbf{A}}$ is related to the section of a linear subspace with the critical locus of a linear projection. We prove that there exists a non-empty Zariski-open set $O \subset GL_n(\mathbb{C})$ such that for all $\mathbf{A} \in O \cap GL_n(\mathbb{Q})$, $f(x)$ is non-negative for all $x \in V \cap \mathbb{R}^n$ if, and only if, $f^{\mathbf{A}}$ can be expressed as a sum of squares of polynomials on the truncated variety generated by the ideal $\langle \mathsf{M}_i^{\mathbf{A}} \rangle$, for $0 \leq i \leq d$. Hence, we can obtain algebraic certificates for lower bounds on $f^\star$ using semidefinite programs. Some numerical experiments are given. We also discuss how to decrease the number of polynomials in $\mathsf{M}_i^{\mathbf{A}}$.

6.17. Gröbner bases and critical points

We consider the problem of computing critical points of the restriction of a polynomial map to an algebraic variety. This is of first importance since the global minimum of such a map is reached at a critical point. Thus, these points appear naturally in non-convex polynomial optimization which occurs in a wide range of scientific applications (control theory, chemistry, economics,etc.). Critical points also play a central role in recent algorithms of effective real algebraic geometry. Experimentally, it has been observed that Gröbner basis algorithms are efficient to compute such points. Therefore, recent software based on the so-called Critical Point Method are built on Gröbner bases engines. Let $f_1, ..., f_p$ be polynomials in $\mathbb{Q}[x_1, ..., x_n]$ of degree D, $V \subset \mathbb{C}^n$ be their complex variety and π_1 be the projection map $(x_1, ..., x_n) \mapsto x_1$. The critical points of the restriction of π_1 to V are defined by the vanishing of $f_1, ..., f_p$ and some maximal minors of the Jacobian matrix Indus associated to $f_1, ..., f_p$. Such a system is algebraically structured: the ideal it generates is the sum of a determinantal ideal and the ideal generated by $f_1, ..., f_p$. In [26], we provide the first complexity estimates on the computation of Gröbner bases of such systems defining critical points. We prove that under genericity assumptions on $f_1, ..., f_p$, the complexity is polynomial in the generic number of critical points, i.e.$D^p(D-1)^{n-p}\begin{pmatrix} n-1 \\ p-1 \end{pmatrix}$ More particularly, in the quadratic case $D = 2$, the complexity of such a Gröbner basis computation is polynomial in the number of variables n and exponential in p. We also give experimental evidence supporting these theoretical results.

PROSECCO Project-Team

6. New Results

6.1. Verification of Security Protocols in the Symbolic Model

The symbolic model of protocols, or Dolev-Yao model is an abstract model in which messages are represented by terms. Our protocol verifier ProVerif relies on this model. This year, we have mainly worked on the verification of protocols with lists and on an extension of ProVerif to prove more observational equivalences.

6.1.1. Verification of Protocols with Lists

Participants: Bruno Blanchet [correspondant], Miriam Paiola.

security protocols, symbolic model, automatic verification, Horn clauses, secrecy

We have designed a novel, simple technique for proving secrecy properties for security protocols that manipulate lists of unbounded length, for an unbounded number of sessions [32]. More specifically, our technique relies on the Horn clause approach used in the automatic verifier ProVerif: we show that if a protocol is proven secure by our technique with lists of length one, then it is secure for lists of unbounded length. Interestingly, this theorem relies on approximations made by our verification technique: in general, secrecy for lists of length one does not imply secrecy for lists of unbounded length. Our result can be used in particular to prove secrecy properties for group protocols with an unbounded number of participants and for some XML protocols (web services) with ProVerif.

6.1.2. Proving More Process Equivalences with ProVerif

Participants: Bruno Blanchet [correspondant], Vincent Cheval.

security protocols, symbolic model, automatic verification, observational equivalence, privacy

We have extended the automatic protocol verifier ProVerif in order to prove more observational equivalences [28]. ProVerif can prove observational equivalence between processes that have the same structure but differ by the messages they contain. In order to extend the class of equivalences that ProVerif handles, we extend the language of terms by defining more functions (destructors) by rewrite rules. In particular, we allow rewrite rules with inequalities as side-conditions, so that we can express tests "if then else" inside terms. Finally, we provide an automatic procedure that translates a process into an equivalent process that performs as many actions as possible inside terms, to allow ProVerif to prove the desired equivalence. These extensions have been implemented in ProVerif and allow us to automatically prove anonymity in the private authentication protocol by Abadi and Fournet.

6.2. Verification of Security Protocols in the Computational Model

The computational model of protocols considers messages as bitstrings, which is more realistic than the formal model, but also makes the proofs more difficult. Our verifier CryptoVerif is sound in this model. This year, we have worked on a compiler from CryptoVerif speficications to OCaml, and we have used CryptoVerif to verify the password-based protocol One-Encryption Key Exchange (OEKE).

6.2.1. Generation of Implementations Proved Secure in the Computational model

Participants: Bruno Blanchet [correspondant], David Cadé.

security protocols, computational model, implementation, verification, compiler

We have designed a novel approach for proving specifications of security protocols in the computational model and generating runnable implementations from such proved specifications. We rely on the computationally-sound protocol verifier CRYPTOVERIF for proving the specification, and we have implemented a compiler that translates a CRYPTOVERIF specification into an implementation in OCaml [26]. We have also proved that this compiler preserves security [27]. We have applied this compiler to the SSH Transport Layer protocol: we proved the authentication of the server and the secrecy of the session keys in this protocol and verified that the generated implementation successfully interacts with OpenSSH. The secrecy of messages sent over the SSH tunnel cannot be proved due to known weaknesses in SSH with CBC-mode encryption.

6.2.2. Proof of One-Encryption Key Exchange using CryptoVerif

Participant: Bruno Blanchet [correspondant].

security protocols, computational model, automatic proofs, formal methods, password-based authentication

We have obtained a mechanized proof of the password-based protocol One-Encryption Key Exchange (OEKE) using the computationally-sound protocol prover CRYPTOVERIF [25]. OEKE is a non-trivial protocol, and thus mechanizing its proof provides additional confidence that it is correct. This case study was also an opportunity to implement several important extensions of CRYPTOVERIF, useful for proving many other protocols. We have indeed extended CRYPTOVERIF to support the computational Diffie-Hellman assumption. We have also added support for proofs that rely on Shoup's lemma and additional game transformations. In particular, it is now possible to insert case distinctions manually and to merge cases that no longer need to be distinguished. Eventually, some improvements have been added on the computation of the probability bounds for attacks, providing better reductions. In particular, we improve over the standard computation of probabilities when Shoup's lemma is used, which allows us to improve the bound given in a previous manual proof of OEKE, and to show that the adversary can test at most one password per session of the protocol.

6.3. New Attacks on RSA PKCS#1 v1.5

Participants: Graham Steel [correspondant], Romain Bardou.

cryptographic hardware, security API, key management, vulnerabilities

RSA PKCS#1v1.5 is the most commonly used standard for public key encryption, used for example in TLS/SSL. It has been known to be vulnerable to a so-called padding-oracle attack since 1998 when Bleichenbacher described the vulnerability at CRYPTO. The attack, known as the "million message attack" was not thought to present a practical threat, due in part to the large number of oracle messages required. In a paper published at CRYPTO 2012 [22] we gave original modifications showing how the attack can be completed in a median of just 15 000 messages. The results lead to widespread interest, indicated by over 1400 downloads of the long version of the paper from the HAL webpage and articles in the New York Times, Boston Globe and Süddeutscher Zeitung.

6.4. Security Proofs for Revocation

Participants: Graham Steel [correspondant], Véronique Cortier, Cyrille Wiedling.

security API, key management, formal methods, security proofs

Revocation of expired or corrupted keys is a common feature of industrially deployed key management systems but an aspect that is almost always missing from formal models. We succeeded in adding revocation to a formal specification of a key management API allowing the proof of strong security properties after corrupted keys are revoked. In particular we showed a self-healing property whereby after a corrupted key expires, after a certain amount of time, the system is safe again. The work was published at ACM CCS 2012.

6.5. Discovering Concrete Attacks on Web Applications by Formal Analysis

Participants: Karthikeyan Bhargavan [correspondant], Sergio Maffeis, Chetan Bansal, Antoine Delignat-Lavaud.

web application security, formal methods, automated verification, vulnerabilities Social sign-on and social sharing are becoming an ever more popular feature of web applications. This success is largely due to the APIs and support offered by prominent social networks, such as Facebook, Twitter, and Google, on the basis of new open standards such as the OAuth 2.0 authorization protocol. A formal analysis of these protocols must account for malicious websites and common web application vulnerabilities, such as cross-site request forgery and open redirectors. We model several configurations of the OAuth 2.0 protocol in the applied pi-calculus and verify them using ProVerif. Our models rely on WebSpi, a new library for modeling web applications and web-based attackers that is designed to help discover concrete website attacks. Our approach is validated by finding dozens of previously unknown vulnerabilities in popular websites such as Yahoo and WordPress, when they connect to social networks such as Twitter and Facebook. This work was published in CSF'12 [21].

To protect sensitive user data against server-side attacks, a number of security-conscious web applications have turned to client-side encryption, where only encrypted user data is ever stored in the cloud. We formally investigate the security of a number of such applications, including password managers, cloud storage providers, an e-voting website and a conference management system. We show that their security relies on both their use of cryptography and the way it combines with common web security mechanisms as implemented in the browser. We model these applications using the WebSpi web security library for ProVerif, we discuss novel attacks found by automated formal analysis, and we propose robust countermeasures. Some of the attacks we discovered were presented at WOOT'12 [24]. Our formal models and verified countermeasures are going to be presented at POST'13 [20].

6.6. Attacks and Proofs for TLS Implementations

Participants: Alfredo Pironti [correspondant], Karthikeyan Bhargavan, Pierre-Yves Strub, Cedric Fournet, Markulf Kohlweiss.

cryptographic protocol, formal methods, automated verification, traffic analysis, vulnerabilities

TLS is possibly the most used secure communications protocol, with a 18-year history of flaws and fixes, ranging from its protocol logic to its cryptographic design, and from the Internet standard to its diverse implementations. We have been engaged in a long-term project on verifying TLS implementations and this project is now coming to fruition, with a number of papers are now in the pipeline. We list two new results below, both are submitted for review.

Websites commonly use HTTPS to protect their users' private data from network-based attackers. By combining public social network profiles with TLS traffic analysis, we present a new attack that reveals the precise identities of users accessing major websites. As a countermeasure, we propose a novel length-hiding scheme that leverages standard TLS padding to enforce website-specific privacy policies. We present several implementations of this scheme, notably a patch for GnuTLS that offers a rich length-hiding API and an Apache module that uses this API to enforce an anonymity policy for sensitive user files. Our implementations are the first to fully exercise the length-hiding features of TLS and our work uncovers hidden timing assumptions in recent formal proofs of these features. Compared to previous work, we offer the first countermeasure that is standards-based, provably secure, and experimentally effective, yet pragmatic, offering websites a precise trade-off between user privacy and bandwidth efficiency. This work is available as an Inria technical report [36].

We develop a verified reference implementation of TLS 1.2. Our code fully supports its wire formats, ciphersuites, sessions and connections, re-handshakes and resumptions, alerts and errors, and data fragmentation, as prescribed in the RFCs; it interoperates with mainstream web browsers and servers. At the same time, our code is carefully structured to enable its modular, automated verification, from its main API down to computational assumptions on its cryptographic algorithms. Our implementation is written in F# and specified in F7. We present security specifications for its main components, such as authenticated stream encryption for the record layer and key establishment for the handshake. We describe their verification using the F7 refinement typechecker. To this end, we equip each cryptographic primitive and construction of TLS with a new typed interface that captures its security prop- erties, and we gradually replace concrete implementations

with ideal functionalities. We finally typecheck the protocol state machine, and thus obtain precise security theorems for TLS, as it is implemented and deployed. We also revisit classic attacks and report a few new ones. This work is under review and will be released as an Inria technical report in January 2013.

SECRET Project-Team

5. New Results

5.1. Symmetric cryptosystems

Participants: Christina Boura, Baudoin Collard, Anne Canteaut, Pascale Charpin, Gohar Kyureghyan, 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 as 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 restricting 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 of 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 secure hash functions. Most teams in the research community are actually working on the design and on the analysis (cryptanalysis and optimisation of the performance) of such primitives.

5.1.1. Hash functions.

Following the recent attacks against almost all existing hash functions (MD5, SHA-0, SHA-1...), we have initiated a research work in this area, especially within the Saphir-2 ANR project and with several PhD theses. Our work on hash functions is two-fold: we have designed two new hash functions, named FSB and Shabal, which have been submitted to the SHA-3 competition, and we have investigated the security of several hash functions, including the previous standards (SHA-0, SHA-1...) and some other SHA-3 candidates.

Recent results:

- Upper bounds on the degree of an iterated permutation from the degree of the inverse of the inner transformation; this result has been applied both to hash functions and to block ciphers. Most notably, this work leads to the best (theoretical) analysis of the hash function Keccak, which has been selected for the new SHA-3 standard [12], [22], [9].
- Side-channel attacks on two SHA-3 candidates, Skein and Grøstl, when they are used with HMAC, and counter-measures [23], [50].
- Indifferentiability results for a broadened mode of operation including the modes based on block ciphers, and modes based on un-keyed functions [51].

5.1.2. Block ciphers.

Even if the security of the current block cipher standard, AES, is not threaten 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 analysed. Most of our work in this area is related to an ANR Project named BLOC.

Recent results:

- Algebraic analysis of some recent lightweight block ciphers, including LED and Piccolo [24].
- Analysis of the security of the lightweight block cipher mCRYPTON [56].
- Design of a new block cipher, named PRINCE, with a very low-latency, leading to instantaneous encryption (i.e., within one clock cycle) with a very competitive chip area [21], [49].
- Analysis of the differential properties of the AES Superbox [58].
- Study of the significance of the related-key and known-key models for block ciphers [48].

5.1.3. Stream ciphers.

The project-team has been involved in the international project eSTREAM, which aimed at recommending some secure stream ciphers.

Recent results:

- Generalisation of several improvements of the so-called correlation attacks against stream ciphers and study of their complexities [13].
- Study of the bias of parity-check relations for combination generators used in stream ciphers [14].

5.1.4. Cryptographic properties and construction of appropriate building blocks.

The construction of building blocks which guarantee a high resistance to 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 the origin of new attacks. This work involves fundamental aspects related to discrete mathematics, cryptanalysis and implementation aspects. Actually, characterising 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.

Recent results:

- Study of the algebraic properties (e.g. the algebraic degree) of the inverses of APN power permutations [26].
- Study of the planarity of some mappings, including products of linearized polynomials [25], [16].
- Definition of a new criterion for Sboxes and link with some recent algebraic attacks on the hash functuion Hamsi [29], [9].
- Survey of PN and APN mappings [42].

5.2. Code-based cryptography

Participants: Grégory Landais, Rafael Misoczki, Nicolas Sendrier, Dimitrios Simos, 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 scheme).

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,
- address new functionalities, like hashing or symmetric encryption.

Recent results:

- A new variant of McEliece using Moderate Density Parity Check (MDPC) codes [55];
- An optimized software implementation of the code-based digital signature scheme CFS [27];
- An attack on a homomorphic encryption scheme [53];
- An attack on a variant of the McEliece cryptosystem based on Reed-Solomon codes [54].

5.3. Error-correcting codes and applications

Participants: Mamdouh Abbara, Marion Bellard, Denise Maurice, Nicolas Sendrier, Jean-Christophe Sibel, Jean-Pierre Tillich, Audrey Tixier.

We mainly investigate two new application domains for decoding algorithms: reverse engineering of communication systems, and quantum error correcting codes for which we have shown that some of them can be decoded successfully with iterative decoding algorithms.

5.3.1. Quantum codes.

The knowledge we have acquired in iterative decoding techniques has also led to study whether or not the very same techniques could also be used to decode quantum codes. Part of the old ACI project “RQ” in which we were involved and the new ANR project “COCQ” are about this topic. It is worth noticing that 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 obtained by transforming a quantum CSS LDPC code into a code over a larger alphabet which improves substantially the performances under iterative decoding [18];
- Construction of spatially coupled quantum LDPC codes which performs well under iterative decoding almost up to the coherent capacity of the quantum channel [19].

5.3.2. Reverse engineering of communication systems.

To evaluate the quality of a cryptographic algorithm, it is usually assumed that its specifications are public, as, in accordance with Kerckhoffs principle [1], 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, is 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 DGA and the French Ministry for Defense.

Recent results:

- Recontruction of the constellation labeling (i.e. used in the modulator of a communication system) in presence of error and when the underlying code is convolutional [20].

[1] Kerckhoffs stated that principle in a paper entitled *La Cryptographie militaire*, published in 1883.

CAD Team

6. New Results

6.1. Geometry Modeling and Processing

6.1.1. Relaxed lightweight assembly retrieval using vector space model

Participants: Kai-Mo Hu, Bin Wang, Jun-Hai Yong, Jean-Claude Paul.

Assembly searching technologies are important for the improvement of design reusability. However, existing methods require that assemblies possess high-level information, and thus cannot be applied in lightweight assemblies. In this paper, we propose a novel relaxed lightweight assembly retrieval approach based on a vector space model (VSM). By decomposing the assemblies represented in a watertight polygon mesh into bags of parts, and considering the queries as a vague specification of a set of parts, the resilient ranking strategy in VSM is successfully applied in the assembly retrieval. Furthermore, we take the scale-sensitive similarities between parts into the evaluation of matching values, and extend the original VSM to a relaxed matching framework. This framework allows users to input any fuzzy queries, is capable of measuring the results quantitatively, and performs well in retrieving assemblies with specified characteristics. To accelerate the online matching procedure, a typical parts based matching process, as well as a greedy strategy based matching algorithm is presented and integrated in the framework, which makes our system achieve interactive performance. We demonstrate the efficiency and effectiveness of our approach through various experiments on the prototype system. [19]

6.1.2. Calculating Jacobian coefficients of primitive constraints with respect to Euler parameters

Participants: Hai-Chuan Song, Jun-Hai Yong.

It is a fundamental problem to calculate Jacobian coefficients of constraint equations in assembly constraint solving because most approaches to solving an assembly constraint system will finally resort to a numerical iterative method that requires the first-order derivatives of the constraint equations. The most-used method of deriving the Jacobian coefficients is to use virtual rotation which is originally presented to derive the equations of motion of constrained mechanical systems. However, when Euler parameters are adopted as the state variables to represent the transformation matrix, using the virtual rotation will yield erroneous formulae of Jacobian coefficients. The reason is that Euler parameters are incompatible with virtual rotation. In this paper, correct formulae of Jacobian coefficients of geometric constraints with respect to Euler parameters are presented in both Cartesian coordinates and relative generalized coordinates. Experimental results show that our proposed formulae make Newton-Raphson iterative method converge faster and more stable. [22]

6.1.3. An extended schema and its production rule-based algorithms for assembly data exchange using IGES

Participants: Kai-Mo Hu, Bin Wang, Jun-Hai Yong.

Assembly data exchange and reuse play an important role in CAD and CAM in shortening the product development cycle. However, current CAD systems cannot transfer mating conditions via neutral file format, and their exported IGES files are heterogeneous. In this paper, a schema for the full data exchange of assemblies is presented based on IGES. We first design algorithms for the pre-and-post processors of parts based on solid model, in which the topologies are explicitly specified and will be referred by mating conditions, and then extend the IGES schema by introducing the Associativity Definition Entity and Associativity Instance Entity defined in IGES standard, so as to represent mating conditions. Finally, a production rule-based method is proposed to analyze and design the data exchange algorithms for assemblies. Within this schema, the heterogeneous representations of assemblies exported from different CAD systems can be processed appropriately, and the mating conditions can be properly exchanged. Experiments on the prototype system verify the robustness, correctness, and flexibility of our schema. [18]

6.1.4. Robust shape normalization of 3D articulated volumetric models

Participants: Yu-Shen Liu, Jun-Hai Yong, Jean-Claude Paul.

3D shape normalization is a common task in various computer graphics and pattern recognition applications. It aims to normalize different objects into a canonical coordinate frame with respect to rigid transformations containing translation, rotation and scaling in order to guarantee a unique representation. However, the conventional normalization approaches do not perform well when dealing with 3D articulated objects.

To address this issue, we introduce a new method for normalizing a 3D articulated object in the volumetric form. We use techniques from robust statistics to guide the classical normalization computation. The key idea is to estimate the initial normalization by using implicit shape representation, which produces a novel articulation insensitive weight function to reduce the influence of articulated deformation. We also propose and prove the articulation insensitivity of implicit shape representation. The final solution is found by means of iteratively reweighted least squares. Our method is robust to articulated deformation without any explicit shape decomposition. The experimental results and some applications are presented for demonstrating the effectiveness of our method. [24]

6.1.5. G^1 continuous approximate curves on NURBS surfaces

Participant: Jun-Hai Yong.

Curves on surfaces play an important role in computer aided geometric design. In this paper, we present a parabola approximation method based on the cubic reparameterization of rational Bézier surfaces, which generates G^1 continuous approximate curves lying completely on the surfaces by using iso-parameter curves of the reparameterized surfaces. The Hausdorff distance between the approximate curve and the exact curve is controlled under the user-specified tolerance. Examples are given to show the performance of our algorithm. [28]

6.1.6. The IFC-based path planning for 3D indoor spaces

Participant: Yu-Shen Liu.

Path planning is a fundamental problem, especially for various AEC applications, such as architectural design, indoor and outdoor navigation, and emergency evacuation. However, the conventional approaches mainly operate path planning on 2D drawings or building layouts by simply considering geometric information, while losing abundant semantic information of building components. To address this issue, this paper introduces a new method to cope with path planning for 3D indoor space through an IFC (Industry Foundation Classes) file as input. As a major data exchange standard for Building Information Modeling (BIM), the IFC standard is capable of restoring both geometric information and rich semantic information of building components to support lifecycle data sharing. The method consists of three main steps: (1) extracting both geometric and semantic information of building components defined within the IFC file, (2) discretizing and mapping the extracted information into a planar grid, (3) and finally finding the shortest path based on the mapping for path planning using Fast Marching Method. The paper aims to process different kinds of building components and their corresponding properties to obtain rich semantic information that can enhance applications of path planning. In addition, the IFC-based distributed data sharing and management is implemented for path planning. The paper also presents some experiments to demonstrate the accuracy, efficiency and adaptability of the method. Video demonstration is available from http://cgcad.thss.tsinghua.edu.cn/liuyushen/ifcpath/. [20]

6.1.7. Recovering Geometric Detail by Octree Normal Maps

Participants: Bin Wang, Jean-Claude Paul.

This paper presents a new approach for constructing normal maps that capture high-frequency geometric detail from dense models of arbitrary topology and are applied to the simplified version of the same models generated by any simplification method to mimic the same level of detail. A variant of loose octree scheme is used to optimally calculate the mesh normals. A B-spline surface fitting based method is employed to solve the issue of thin plate. A memory saving Breadth-First Search (BFS) order construction is designed. Furthermore, a speedup scheme that exploits access coherence is used to accelerate filtering operation. The proposed method can synthesize good quality images of models with extremely high number of polygons while using much less memory and render at much higher frame rate. [31]

6.1.8. An improved example-driven symbol recognition approach in engineering drawings

Participants: Hui Zhang, Ya-Mei Wen.

In this paper, an improved example-driven symbol recognition algorithm is proposed for CAD engineering drawings. First, in order to represent the structure of symbols more clearly and simply, we involve the text entity as one of the basic elements and redefine the relation representation mechanism, which is the foundation for the following algorithms. Then, the structure graph and a constrained tree can be established automatically for the target symbol, using the knowledge acquisition algorithm. In this process, the highest priority element is considered as the key feature, which will be regarded as the root node of the tree. The sequence of breadth first traveling will be recorded to be the recognition rule and saved in the symbol library. In the recognition process, the nodes with the same type as the key features can be located first in the drawing. Unnecessary matching calculations would be greatly reduced because of the accurate location. The other elements around, which satisfy the topology structure of the constrained tree, will be found next. The target symbol is recognized if all of the elements and constraints in the tree are found. Moreover, an extra preprocessing analysis approach is proposed to address repeat modes in a symbol. Thus, similar symbols can be recognized by one rule. We evaluate the proposed approach on the GREC databases and the real engineering drawings. The experimental results validate its effectiveness and efficiency. [17]

6.1.9. 3DMolNavi: A web-based retrieval and navigation tool for flexible molecular shape comparison

Participants: Yu-Shen Liu, Jean-Claude Paul.

Many molecules of interest are flexible and undergo significant shape deformation as part of their function, but most existing methods of molecular shape comparison treat them as rigid shapes, which may lead to incorrect measure of the shape similarity of flexible molecules. Currently, there still is a limited effort in retrieval and navigation for flexible molecular shape comparison, which would improve data retrieval by helping users locate the desirable molecule in a convenient way. To address this issue, we develop a web-based retrieval and navigation tool, named 3DMolNavi, for flexible molecular shape comparison. This tool is based on the histogram of Inner Distance Shape Signature (IDSS) for fast retrieving molecules that are similar to a query molecule, and uses dimensionality reduction to navigate the retrieved results in 2D and 3D spaces. We tested 3DMolNavi in the Database of Macromolecular Movements (MolMovDB) and CATH. Compared to other shape descriptors, it achieves good performance and retrieval results for different classes of flexible molecules. The advantages of 3DMolNavi, over other existing softwares, are to integrate retrieval for flexible molecular shape comparison and enhance navigation for user's interaction. [23]

6.1.10. Manifold-ranking based retrieval using k-regular nearest neighbor graph

Participants: Bin Wang, Kai-Mo Hu, Jean-Claude Paul.

Manifold-ranking is a powerful method in semi-supervised learning, and its performance heavily depends on the quality of the constructed graph. In this paper, we propose a novel graph structure named k-regular nearest neighbor (k-RNN) graph as well as its constructing algorithm, and apply the new graph structure in the framework of manifold-ranking based retrieval. We show that the manifold-ranking algorithm based on our proposed graph structure performs better than that of the existing graph structures such as k-nearest neighbor (k-NN) graph and connected graph in image retrieval, 2D data clustering as well as 3D model retrieval. In addition, the automatic sample reweighting and graph updating algorithms are presented for the relevance

feedback of our algorithm. Experiments demonstrate that the proposed algorithm outperforms the state-of-the-art algorithms. [25]

6.2. Computer Graphics

6.2.1. Content-Based Color Transfer

Participants: Fuzhang Wu, Weiming Dong, Yan Kong, Xing Mei, Jean-Claude Paul, Xiaopeng Zhang.

This paper presents a novel content-based method for transferring the colour patterns between images. Unlike previous methods that rely on image colour statistics, our method puts an emphasis on high-level scene content analysis. We first automatically extract the foreground subject areas and background scene layout from the scene. The semantic correspondences of the regions between source and target images are established. In the second step, the source image is re-coloured in a novel optimization framework, which incorporates the extracted content information and the spatial distributions of the target colour styles. A new progressive transfer scheme is proposed to integrate the advantages of both global and local transfer algorithms, as well as avoid the over-segmentation artefact in the result. Experiments show that with a better understanding of the scene contents, our method well preserves the spatial layout, the colour distribution and the visual coherence in the transfer process. As an interesting extension, our method can also be used to re-colour video clips with spatially-varied colour effects. [26]

6.2.2. Large-scale forest rendering: Real-time, realistic, and progressive

Participants: Xiaopeng Zhang, Weiming Dong.

Real-time rendering of large-scale forest landscape scenes is important in many applications, such as video games, Internet graphics, and landscape and cityscape scene design and visualization. One challenge in the field of virtual reality is transferring a large-scale forest environment containing plant models with rich geometric detail through the network and rendering them in real time. We present a new framework for rendering large-scale forest scenes realistically and quickly that integrates extracting level of detail (LOD) tree models, rendering real-time shadows for large-scale forests, and transmitting forest data for network applications. We construct a series of LOD tree models to compress the overall complexity of the forest in view-dependent forest navigation. A new leaf phyllotaxy LOD modeling method is presented to match leaf models with textures, balancing the visual effect and model complexity. To progressively render the scene from coarse to fine, sequences of LOD models are transferred from simple to complex. The forest can be rendered after obtaining a simple model of each tree, allowing users to quickly see a sketch of the scene. To improve client performance, we also adopt a LOD strategy for shadow maps. Smoothing filters are implemented entirely on the graphics processing unit (GPU) to reduce the shadows' aliasing artifacts, which creates a soft shadowing effect. We also present a hardware instancing method to render more levels of LOD models, which overcomes the limitation of the latest GPU that emits primitives into only a limited number of separate vertex streams. Experiments show that large-scale forest scenes can be rendered with smooth shadows and in real time. [14]

6.2.3. Fast Multi-Operator Image Resizing and Evaluation

Participants: Weiming Dong, Xiaopeng Zhang, Jean-Claude Paul.

Current multi-operator image resizing methods succeed in generating impressive results by using image similarity measure to guide the resizing process. An optimal operation path is found in the resizing space. However, their slow resizing speed caused by inefficient computation strategy of the bidirectional patch matching becomes a drawback in practical use. In this paper, we present a novel method to address this problem. By combining seam carving with scaling and cropping, our method can realize content-aware image resizing very fast. We define cost functions combing image energy and dominant color descriptor for all the operators to evaluate the damage to both local image content and global visual effect. Therefore our algorithm can automatically find an optimal sequence of operations to resize the image by using dynamic programming or greedy algorithm. We also extend our algorithm to indirect image resizing which can protect the aspect ratio of the dominant object in an image. [16]

6.2.4. Easy modeling of realistic trees from freehand sketches

Participant: Xiaopeng Zhang.

Creating realistic 3D tree models in a convenient way is a challenge in game design and movie making due to diversification and occlusion of tree structures. Current sketch-based and image-based approaches for fast tree modeling have limitations in effect and speed, and they generally include complex parameter adjustment, which brings difficulties to novices. In this paper, we present a simple method for quickly generating various 3D tree models from freehand sketches without parameter adjustment. On two input images, the user draws strokes representing the main branches and crown silhouettes of a tree. The system automatically produces a 3D tree at high speed. First, two 2D skeletons are built from strokes, and a 3D tree structure resembling the input sketches is built by branch retrieval from the 2D skeletons. Small branches are generated within the sketched 2D crown silhouettes based on self-similarity and angle restriction. This system is demonstrated on a variety of examples. It maintains the main features of a tree: the main branch structure and crown shape, and can be used as a convenient tool for tree simulation and design. [21]

6.2.5. Real-time ink simulation using a grid-particle method

Participants: Shibiao Xu, Xing Mei, Weiming Dong, Xiaopeng Zhang.

This paper presents an effective method to simulate the ink diffusion process in real time that yields realistic visual effects. Our algorithm updates the dynamic ink volume using a hybrid grid-particle method: the fluid velocity field is calculated with a low-resolution grid structure, whereas the highly detailed ink effects are controlled and visualized with the particles. To facilitate user interaction and extend this method, we propose a particle-guided method that allows artists to design the overall states using the coarse-resolution particles and to preview the motion quickly. To treat coupling with solids and other fluids, we update the grid-particle representation with no-penetration boundary conditions and implicit interaction conditions. To treat moving "ink-emitting" objects, we introduce an extra drag-force model to enhance the particle motion effects; this force might not be physically accurate, but it proves effective for producing animations. We also propose an improved ink rendering method that uses particle sprites and motion blurring techniques. The simulation and the rendering processes are efficiently implemented on graphics hardware at interactive frame rates. Compared to traditional fluid simulation methods that treat water and ink as two mixable fluids, our method is simple but effective: it captures various ink effects, such as pinned boundaries and filament patterns, while still running in real time, it allows easy control of the animation, it includes basic solid-fluid interactions, and it can address multiple ink sources without complex interface tracking. Our method is attractive for animation production and art design.

6.2.6. Image zooming using directional cubic convolution interpolation

Participant: Weiming Dong.

Image-zooming is a technique of producing a high-resolution image from its low-resolution counterpart. It is also called image interpolation because it is usually implemented by interpolation. Keys' cubic convolution (CC) interpolation method has become a standard in the image interpolation field, but CC interpolates indiscriminately the missing pixels in the horizontal or vertical direction and typically incurs blurring, blocking, ringing or other artefacts. In this study, the authors propose a novel edge-directed CC interpolation scheme which can adapt to the varying edge structures of images. The authors also give an estimation method of the strong edge for a missing pixel location, which guides the interpolation for the missing pixel. The authors' method can preserve the sharp edges and details of images with notable suppression of the artefacts that usually occur with CC interpolation. The experiment results demonstrate that the authors'method outperforms significantly CC interpolation in terms of both subjective and objective measures. [30]

CLASSIC Project-Team

5. New Results

5.1. Contributions earlier to 2012 but only published in 2012

Participants: Gérard Biau, Vincent Rivoirard, Gilles Stoltz, Olivier Catoni.

We do not discuss here the contributions provided by [16], [17], [11], [13], [14], [15], [18], since they were achieved in 2011 or earlier (but only published this year due to the reviewing and publishing process). Also, the book [25] (whose first edition was published in 2009) was augmented and revised for its second edition, published this year.

5.2. Extended journal versions written in 2012 of conference papers published in 2011

Participants: Sébastien Gerchinovitz, Gilles Stoltz.

We wrote extended journal papers of some conference papers discussed in previous annual activity reports; they correspond to refences [32], [19], [20].

5.3. Bayesian methods

Participants: Gérard Biau, Vincent Rivoirard.

5.3.1. The ABC method

Approximate Bayesian Computation (ABC for short) is a family of computational techniques which offer an almost automated solution in situations where evaluation of the posterior likelihood is computationally prohibitive, or whenever suitable likelihoods are not available. In the paper [29] Gérard Biau and his coauthors analyze the procedure from the point of view of k-nearest neighbor theory and explore the statistical properties of its outputs. They discuss in particular some asymptotic features of the genuine conditional density estimate associated with ABC, which is a new interesting hybrid between a k-nearest neighbor and a kernel method.

5.3.2. Semi-parametric version of the Bernstein-von Mises theorem

In [22], Vincent Rivoirard and Judith Rousseau study the asymptotic posterior distribution of linear functionals of the density by deriving general conditions to obtain a semi-parametric version of the Bernstein-von Mises theorem. The special case of the cumulative distributive function evaluated at a specific point is widely considered. In particular, they show that for infinite dimensional exponential families, under quite general assumptions, the asymptotic posterior distribution of the functional can be either Gaussian or a mixture of Gaussian distributions with different centering points. This illustrates the positive but also the negative phenomena that can occur for the study of Bernstein-von Mises results. In [22] Vincent Rivoirard and Judith Rousseau use convergence rates on Besov spaces established in [23].

5.4. Sequential learning

Participants: Pierre Gaillard, Gilles Stoltz.

5.4.1. Bandit problems

The article [30] revisits asymptotically optimal results of Lai and Robbins, Burnetas and Katehakis in a non-asymptotic way. A preliminary attempt was mentioned in the 2011 annual report; it was concerned (essentially) with the case of Bernoulli distributions over the arms. We achieve here the stated optimality of the regret bounds for larger models: regular exponential families; finitely supported distributions.

5.4.2. Theoretical results for the prediction of arbitrary sequences

We generalize and unify in [24] several notions of regret under a same banner: these include adaptive regret (regret against a fixed convex combination on subintervals of the time); shifting regret (regret against a slowly evolving target sequence of convex combinations); and discounted regret (when the instances are weighted with weights depending on how recent the instances are). We recover and sometimes improve some earlier bounds.

5.4.3. Forecasting of the production data of oil reservoirs

We applied our sequential aggregation techniques to a new data set, with IFP Energies nouvelles as a partner. The goal was to aggregate in a sequential fashion the forecasts made by some (about 100) base experts in order to predict some behaviors (gas/oil ratio, cumulative oil extracted, water cut) of the exploitation of some oil wells. Results were obtained with the help of an intern, Charles-Pierre Astolfi, and are described in the technical report [27] (to be transformed into a regular journal / conference paper next year).

5.5. Regression, classification, regression methods

Participants: Gérard Biau, Olivier Catoni, Ilaria Giulini.

5.5.1. Metric-based decision procedures

We know now that a good part of the statistical performance of regression and classification algorithms relies on the metric chosen to represent the proximity between the data points. Throughout his work, Gérard Biau became convinced that, well beyond the traditional distances, (dis)similarities and other self-reproducing kernel metrics, it is now necessary to attempt to define proximities generated by the sample itself. These metrics are inevitably random and probabilistic, and force us to rethink the nature of the estimates, as shown for example in the preliminary article [12].

5.5.2. Unsupervised classification in reproducing kernel Hilbert spaces

In her PhD started in September 2012, Ilaria Giulini uses dimension free estimates of the principal components of an i.i.d. sample of points in a Reproducing Kernel Hilbert Space to derive new unsupervised clustering algorithms based on the idea of dimension reduction by nonlinear coordinate smoothing along aggregated principal components. The dimension free estimates are obtained using PAC-Bayes bounds derived from thresholded exponential moments.

5.6. Sparsity and ℓ_1–regularization

Participant: Vincent Rivoirard.

5.6.1. For multivariate Hawkes processes

Motivated by statistical problems in neuroscience, Vincent Rivoirard and his coauthors study in [31] non-parametric inference for multivariate Hawkes processes depending on an unknown function to be estimated by linear combinations of a fixed dictionary. To select coefficients, they propose a Lasso-type methodology where data-driven weights of the penalty are derived from new Bernstein-type inequalities for martingales. Oracle inequalities are established under assumptions on the Gram matrix of the dictionary. Non-asymptotic probabilistic results are proven, which allows them to check these assumptions by considering general dictionaries based on histograms, Fourier or wavelet bases. They finally carry out a simulation study and compare their methodology with the adaptive Lasso procedure proposed by Zou. They observe an excellent behavior of their procedure with respect to the problem of supports recovery. Unlike adaptive Lasso of Zou, their tuning procedure is proven to be robust with respect to all the parameters of the problem, revealing its potential for concrete purposes in neuroscience, but also in other fields.

5.6.2. In the spherical convolution model

In [21], Thanh Mai Pham Ngoc and Vincent Rivoirard consider the problem of estimating a density of probability from indirect data in the spherical convolution model. They aim at building an estimate of the unknown density as a linear combination of functions of an overcomplete dictionary. The procedure is devised through a well-calibrated ℓ_1–penalized criterion. The dictionary approach allows to combine various bases and thus enhances estimates sparsity. They provide an oracle inequality under global coherence assumptions. Moreover, the calibrated procedure that they put forward gives very satisfactory results in the numerical study when compared with other procedures.

5.6.3. For semiparametric nonlinear mixed-effects models

Semiparametric nonlinear mixed-effects models (SNMMs) have been proposed as an extension of nonlinear mixed-effects models (NLMMs). These models are a good compromise and retain nice features of both parametric and nonparametric models resulting in more flexible models than standard parametric NLMMs. In [28], Vincent Rivoirard and his coauthors propose new estimation strategies in SNMMs. They propose a Lasso-type method to estimate the unknown nonlinear function. They derive oracle inequalities for this nonparametric estimator. They combine the two approaches in a general estimation procedure that they illustrate with simulations and through the analysis of a real data set of price evolution in on-line auctions.

5.7. Computational linguistics

Participants: Olivier Catoni, Thomas Mainguy.

In a forthcoming paper, Olivier Catoni and Thomas Mainguy study a new statistical model to learn the syntactic structure of natural languages from a training set made of written sentences. This model learns a new type of stochastic grammar and defines a statistical model on sentences. Global constraints are enforced, that set the approach apart from the family of Markov models. On the other hand, the grammar model generates outputs through a split and merge stochastic process that is more elaborate than the production rules defining a context free grammar. Experiments made on small corpora are very encouraging. Working on large corpora will require to speed up the algorithms used to implement the model as well as some code optimization.

GAMMA3 Project-Team

3. New Results

3.1. Validité des éléments finis usuels

Participants: Houman Borouchaki, Paul-Louis George [correspondant].

éléments finis-éléments finis généralisés-P1-P2-Q1-Q2-Bézier

On continue l'étude sur les conditions assurant la validité géométrique des éléments finis usuels de degré 1 et 2. La formulation éléments finis ne conduisant pas toujours à une conclusion simple, on formule les éléments finis sous leur forme de Bézier. Ceci conduit à exhiber des conditions suffisantes (parfois nécessaires et suffisantes) de validité des éléments, c'est-à-dire des conditions garantissant la positiivité de leur jacobien. Pour les éléments de degré 2, on donne l'interprétion géométrique de ces conditions. Les éléments étudiés sont le triangle à 3 nœuds, le triangle à 6 nœuds, le quadrilatère à 4 nœuds et les quadrilatère à 8 et 9 nœuds, le tétraèdre à 4 nœuds et le tétraèdre à 10 nœuds puis les pentaèdres à 6, 15 et 18 nœuds et les hexaèdres à 8, 27 et 20 nœuds.

On regarde ensuite les éléments finis généralisés déduits d'une formulation en Bézier rationnels puis basés sur des fonctions B-splines et Nurbs.

3.2. Maillages tétraédriques de grande taille

Participants: Houman Borouchaki, Paul-Louis George [correspondant], Loïc Maréchal.

Triangulation-tétraèdre p1-Hilbert- Maillage de grande taille

Le comportement en complexité des algorithmes de triangulation sur les "gros" maillage nous amène à utiliser les algorithmes de renumérotation de type Hilbert qui minimisent les défauts de cache. Cette technique est également utilisée comme aide à l'optimisation des "gros" maillages avec des gains en temps important. L'algorithme de renumérotation est multi-cœurs.

Des triangulations de plusieurs dizaines de millions de sommets sont construites en utilisant un "simple" ordinateur. La vitesse d'insertion frole le million de tétraèdres à la seconde.

Par coquetterie (et pour améliorer la robustesse dans l'absolu), on regarde ce que donne nos méthodes quand on construit des maillages de plus de un milliard de tétraèdres en séquentiel (une machine de un Tera de mémoire est utilisée). On vérifie que la taille des cavités peut être arbitrairement grande ce qui nécessite une programmation plus délicate permettant de traiter ces cas peu courants dans les situations habituelles.

3.3. Surface meshing with metric gradation control

Participants: Patrick Laug [correspondant], Houman Borouchaki.

Scientific computing requires the automatic generation of high quality meshes, in particular isotropic or anisotropic meshes of surfaces defined by a CAD modeler. For this purpose, two major approaches are called direct and indirect. Direct methods (octree, advancing-front or paving) work directly in the tridimensional space, while indirect methods consist in meshing each parametric domain and mapping the resulting mesh onto the composite surface. Using the latter approach, we propose a general scheme for generating “geometric” (or geometry-preserving) meshes by means of metrics. In addition, we introduce a new methodology for controlling the metric gradation in order to improve the shape quality. Application examples have shown the capabilities of this approach.

3.4. Metric field interpolation

Participants: Patrick Laug [correspondant], Houman Borouchaki.

To solve a physical problem formulated in terms of partial differential equations, the finite element method is generally used, based on a spatial discretization, or *mesh*, of the domain studied. Local adaptations of meshes to the behavior of the physical phenomena can improve the accuracy to the computed solutions, and in particular it is possible to capture high variations of the solution in specific areas while maintaining a reasonable number of degrees of freedom. In an initial phase, a mesh of the domain is built by using any particular method, then a first calculation of the solution of the problem is made. After choosing an appropriate criterion (Hessian and/or gradient of the solution, error estimate in general), areas that must be adapted by refinement or coarsening are detected in the initial mesh, and a new mesh is generated which is better adapted to the problem. This process is iterated until obtaining a mesh which satisfies the specified criterion (for which the finite element error is bounded by a specified threshold).

In practice, via an *a posteriori* analysis of the finite element error, a discrete map of sizes or metrics is set to the mesh vertices. This discrete size or metric field is made continuous by interpolating on the mesh, and the new mesh is generated according to this new field. In general, for a given point of the domain, a mesh element containing this point is found, and the interpolation of the size or metric field at this point is made from the sizes or metrics associated with the vertices of the containing element. For a scalar size field, the interpolation is straightforward by considering any interpolation scheme (for instance linear or geometric). On the other hand, the same scheme cannot be applied in the case of metrics representing a tensor field. However, several approaches have been proposed based on the link between a size and the corresponding metric and, in most cases, the interpolation scheme for sizes is applied to a power or the logarithm of the metrics. In particular, as a size h is represented by the isotropic metric $\mathcal{M} = \frac{1}{h^2}\,\mathcal{I}$, where $\mathcal{I}$ is the identity matrix, a possible link consists in approximating the size by $\mathcal{M}^{-\frac{1}{2}}$, then applying the size interpolation scheme to this new metric and finally recovering the interpolated metric. These schemes are still an approximation and require the calculation of the eigenvalues of $\mathcal{M}$ which is generally costly.

In this work, a new method for interpolating discrete metric fields is proposed. It is based on the "natural decomposition" of metrics using the LU factorization. With this decomposition, for each metric, the natural sizes along particular (or natural) directions can be retrieved, thus the size interpolation scheme can be applied to both natural directions and sizes, and the interpolation on the metrics is obtained. The proposed method is faster than those mentioned above and provides a continuous metric field with low variations. Some numerical examples illustrate our methodology.

3.5. Large deformation simulation using adaptive remeshing

Participants: Patrick Laug [correspondant], Houman Borouchaki.

The object of non-linear solid and structural mechanics is the modeling and the computation of structures with strong non-linearities, both geometrical and physical. The aim is to simulate the behavior of a mechanical part submitted to various mechanical stresses, in order to improve its mechanical strength, or even to optimize its manufacturing process with respect to damage occurrence. Among various theoretical, numerical and geometric tools involved in such a simulation, the interest in adaptive remeshing is really high nowadays. It is generally based on local refinement (governed by error estimation) and vertex smoothing strategies. Let us mention that the main difficulty lies in the fact that, in large strains, the domain geometry is variable and cannot be defined in an explicit way.

New contributions to the strategy using adaptive meshing and *a posteriori* error estimation in large elasto-plasticity have been developed. We are interested in the problem of remeshing a mechanical structure composed of several parts (which are in contact) subjected to large plastic deformations. A general scheme, constituted by several steps necessary to an almost optimal representation of the evolving domain, is proposed. These steps are divided into two main categories: the definition of the boundary of the deformed parts and the whole remeshing of the parts. The remeshing is governed by a mesh size map representing the conformity with the underlying geometry of the deformed parts, the improvement of the accuracy of the desired mechanical fields, and the convergence of the mechanical process as well. This size map results from an *a posteriori* estimation of the "interpolation error" independently from the considered mechanical fields. The final deformation after the whole simulation is assumed to be obtained iteratively by "small" deformations

(which is the case in the framework of an explicit integration scheme to solve the problem). After such a small deformation, rigid parts are moved and deformable parts are slightly distorted (assuming that each mesh element is still valid). The remeshing is applied to deformable parts after each deformation increment. The proposed technique is used to simulate the impact of a projectile on a confined explosive. We show in particular that the ignition of the explosive appears in two different areas.

3.6. Maillage d'un milieu géologique et d'ouvrages de stockage

Participants: Patrick Laug [correspondant], Houman Borouchaki.

Cette étude a été menée dans le cadre du partenariat stratégique ANDRA/Inria. L'objectif est la construction d'un maillage statique 3D prenant en compte la géométrie des couches d'un milieu géologique et celle d'ouvrages de stockage afin de réaliser un calcul d'hydraulique et de transfert de solutés. En particulier, ce maillage sera exploité pour mener des calculs préparatoires aux calculs de sûreté. Il permettra de mieux représenter à l'échelle du milieu géologique les différentes voies de transfert (ouvrages et géologie multi-couches) des radionucléides, en considérant les évolutions géodynamiques, et de contribuer à identifier les simplifications éventuelles qui seront définies pour établir le modèle conceptuel de calcul de performances et de sûreté.

Les données d'entrée représentent la description géométrique du milieu géologique incluant les ouvrages de stockage. Le schéma de construction comprend quatre étapes :

1. Prétraitement des données d'entrée. Les sommets multiples du maillage volumique sont fusionnés afin de pouvoir extraire une topologie conforme. Grâce à cette topologie, les surfaces interfaces entre deux couches consécutives sont identifiées. Ces surfaces représentent des contraintes surfaciques que le mailleur volumique doit respecter. En outre, les lignes intersections entre ces surfaces contraintes, appelées lignes d'affleurement, sont identifiées. De même, ces lignes représentent des contraintes linéiques pour le mailleur volumique. Afin de définir la ligne polygonale associée à chaque rivière, les arêtes de l'enveloppe supérieure du maillage volumique de référence (surface topographique) dont les deux extrémités ont le même code de rivière sont identifiées.

2. Définition de la géométrie du domaine 2D de référence. On définit le plan de référence comme étant le plan d'équation $z = 0$, et le domaine 2D de référence comme la trace du polygone de l'extension horizontale dans ce plan. Toutes les contraintes linéiques (lignes d'affleurement, rivières et contours des ouvrages) sont projetées verticalement sur le plan de référence et leurs traces dans le domaine de référence sont retenues. En outre, des nouvelles lignes contraintes parallèles aux contours des ouvrages sont insérées afin de mieux contrôler la génération du maillage du domaine de référence. L'ensemble de toutes les lignes du domaine de référence est rendu conforme par ajout des points aux intersections éventuelles de ces lignes, et aussi par fusion des points et des lignes coïncidents.

3. Construction du maillage quad-dominant du domaine 2D de référence. Le maillage du domaine de référence est généré en utilisant un schéma adaptatif de construction de maillages quad-dominants. Dans un premier temps, un maillage quad-dominant initial du domaine est construit en spécifiant une taille fixe sur les lignes d'affleurement et les rivières et une taille dépendant de la grandeur des ouvrages sur ces derniers. Afin de contrôler la gradation du maillage (rapport maximal entre les longueurs d'arêtes issues d'un même sommet), deux maillages quad-dominants adaptés sont générés. Ici, l'adaptation consiste à modifier la carte de taille courante pour respecter le seuil de gradation spécifié.

4. Construction du maillage hex-dominant 3D du milieu. Le maillage volumique du milieu géologique est généré par extrusion verticale du maillage quad-dominant du domaine de référence. Deux types de configuration sont considérés : extrusion d'un quadrilatère (dit de base) du maillage du domaine de référence et extrusion d'un triangle (dit de base) du maillage du domaine de référence. Dans le premier cas, selon la configuration des surfaces (surfaces interfaces entre deux couches ou faces supérieures ou inférieures d'ouvrages) rencontrées, des hexaèdres et des prismes sont générés. Plus précisément, dans ce cas, l'extrusion résulte en un ensemble de quadrilatères ordonnés verticalement avec quatre arêtes appartenant à la même surface ou deux arêtes opposées appartenant chacune à une surface. Les quadrilatères consécutifs sont

connectés et, en fonction du nombre de sommets communs entre deux quadrilatères consécutifs, des hexaèdres ou des prismes sont générés. Par ailleurs, une configuration de quadrilatère est validée si d'une part chaque élément résultant est géométriquement valide (hexaèdre, prisme ou pyramide) et si, d'autre part, il contient son barycentre et ses faces sont quasi-planes. Dans le cas contraire, le quadrilatère de base est subdivisé en deux triangles et généralement selon la diagonale donnant une configuration de deux triangles de Delaunay.

3.7. Advanced meshing and remeshing procedure for mechanical and numerical simulations

Participants: Abel Cherouat [correspondant], Houman Borouchaki, Paul-Louis George, Patrick Laug, Zhu Aichun, Jie Zhang, Faouzi Slimani, Guillaume Dufaye.

Most metal forming parts involve complex geometry and flow characteristics as large (visco)-plasticity flow, heat exchange, ductile damage, evolving contact with friction. An intrinsic difficulty in metal forming process is the constantly changing configuration of the deforming part (finite transformation, thermo-plastic flow). In metal forming, the mesh size should be adapted to the curvature of complex tools in order to optimize the contact boundaries and the damaged zones. These problems can be resolved if an adaptive remeshing scheme is incorporated automatically in the finite element analysis. It is necessary to adapt the mesh in order to improve the geometry of the deformed part and the damage localization. To mesh the 3D computational domain, we apply a new optimization approach which uses a combined Delaunay-frontal method to define field points and to construct the connection between these points or with a given prescribed size map (error estimate). The first objective of this project is to develop a 3D advanced remeshing procedure (error estimation, field transfer, optimisation meshing) for metal forming. The second objective is to integrate in a computational environment the mechanical model, 3D reconstruction from images, reliability-optimisation and the remeshing procedure using the ABAQUS/Explicit solver and the adaptive mesher. Application is dedicated to some examples (side pressing, blanking and orthogonal cutting, 3D guillotining, thermo-hydroforming and forging) for metal forming and breast and porous metal foam material reconstitution.

3.8. Effect of fibre geometrical morphology on the mechanical properties of PolyPropylene Hemp fibre composite material

Participants: Abel Cherouat [correspondant], Florent Ilczyszyn.

These last years, hemp fibres have been used as reinforcement for compound based on polymer in different industrial manufacturing for their interesting mechanical and ecological properties. Hemp fibres present a non-homogeneous cross section and complex geometry that can have a high effect on their mechanical properties. The mechanical properties of hemp fibres are rather difficult to determine and request a specific characterization method. In this project, micro-tensile tests coupled with numerical imaging treatments, meshing reconstitution and finite elements computations are investigated. The numerical imaging allows to define finely the hemp cross section along the fibre and aims to reconstruct a 3D hemp fibre CAD using adaptive mesh.

3.9. Mise au point de méthodes de remaillage adaptatif 3D dans le cadre de simulations numériques de mise en forme de structure minces

Participants: Houman Borouchaki, Abel Cherouat, Laurence Moreau [correspondant].

Au cours des simulations numériques de mise en forme en 3D, les grandes déformations mises en jeu font que le maillage subit de fortes distorsions. Il est alors nécessaire de remailler continuellement la pièce afin de pouvoir capturer les détails géométriques des surface en contact, adapter la taille du maillage à la solution physique et surtout pouvoir effectuer la simulation jusqu'à la fin du procédé de mise en forme. Lorsque la pièce est comprise entre des outils rigides (cas de l'emboutissage), aux problèmes de remaillage s'ajoutent aussi des difficultés sur la gestion du contact entre les pièce. Une méthode couplant une stratégie de remaillage adaptatif et une technique de projection a été développée. La méthode de remaillage adaptatif, basée sur des techniques

de raffinement et déraffinement est contrôlée par des cartes de taille géométrique et physique. La projection des nouveaux nœuds sur l'outil permet de conserver le contact entre la pièce et l'outil. Afin de pouvoir réaliser des simulations numériques de composites tissés, une procédure spécifique a été ajoutée au remailleur afin de pouvoir raffiner les éléments finis bi-composants (association d'éléments finis de barre et de membrane orientés matérialisant le comportement de fibres chaîne et trame). Le formage incrémental est un procédé de mise en forme de tôle récent sans poinçon ni matrice, basé sur la déformation progressive du flan à l'aide d'un simple outil de forme hémisphérique commandé par une machine à commande numérique. L'inconvénient de ce nouveau procédé étant le temps de calcul, nous avons proposé une méthode de remaillage adaptatif permettant de raffiner le maillage uniquement au voisinage de l'outil rigide, là où les déformations ont lieux et permettant de déraffiner le maillage après le passage de l'outil rigide.

3.10. Mise au point de méthodes de remaillage adaptatif 3D dans le cadre de simulations numériques de mise en forme de structure minces

Participants: Houman Borouchaki, Abel Cherouat, Laurence Moreau [correspondant].

L'objectif est de reconstruire un maillage de la surface 3D d'un buste féminin à partir d'images 2D issues des prises de vue simultanées de plusieurs appareils photos numériques (photos prises sous des angles différents). Une cabine de mesure, équipée de 24 appareils photos numériques, 6 vidéoprojecteurs, pilotée par un ordinateur extérieur à la cabine a été développée et permet d'acquérir de manière simultanée 24 photos numériques du buste sous des angles différents. Un algorithme original basé sur l'utilisation d'un motif projeté sur le buste a été développé et programmé pour la corrélation entre les images 2D. Une méthode de triangulation 3D associée à une technique d'optimisation a été développée et permet de déterminer les positions 3D des points à partir des pixels de vues différentes.

3.11. Applications du maillage et développements de méthodes avancées pour la cryptographie

Participants: Dominique Barchiesi [correspondant], Thomas Grosges, 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.

3.12. Développement de méthodes avancées et maillages appliqués àa l'étude de la nanomorphologie des nanotubes/fils en suspension liquide

Participants: Dominique Barchiesi, Houman Borouchaki, Abel Cherouat, Anis Chaari, Thomas Grosges [correspondant], Laurence Moreau.

Ce projet de recherche (NANOMORPH) a pour objet principal le de´veloppement et la mise au point d'une instrumentation optique pour de´terminer la distribution en tailles et le coefficient de forme de nanofils (NF) ou de nanotubes (NT) en suspension dans un e´coulement. Au cours de ce projet, deux types de techniques optiques comple´mentaires sont de´veloppe´es. La premie`re, base´e sur la diffusion statique de la lumie`re, ne´cessite d'e´tudier au pre´alable la physico-chimie de la dispersion, la stabilisation et l'orientation des nanofils dans les milieux d'e´tude. La seconde me´thode, base´e sur une me´thode opto-photothermique pulse´e, ne´cessite en sus, la mode´lisation de l'interaction laser/nanofils, ainsi que l'e´tude des phe´nome`nes multiphysiques induits par ce pro-cessus. L'implication de l'e´quipe-projet GAMMA3 concerne principalement la simulation mul- tiphysique de l'interaction laser-nanofils et l'e´volution temporelle des bulles et leurs formations. L'une des principales difficulte´s de ces proble´matiques est que la ge´ome´trie du domaine est variable (a` la fois au sens ge´ome´trique et topologique). Ces simulations ne peuvent donc eˆtre re´alise´es que dans un sche´ma adaptatif de calcul ne´cessitant le remaillage tridimensionnel mobile, de´formable avec topologie variable du domaine (formation et e´volution des bulles au cours du temps et de l'espace).

3.13. 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], Thomas Grosges, 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 une 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 2012 a constitué en une pré-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.

3.14. Mesh adaptation for very high-order numerical scheme

Participants: Frederic Alauzet [correspondant], Adrien Loseille, Estelle Mbinky.

In the past, we have demonstrate that multi-scale anisotropic mesh adaptation is a powerful tool to accurately simulate compressible flow problem and to obtain faster convergence to continuous solutions. But, this was limited to second order numerical scheme. Nowadays, numerous teams are working on the development of very high-order numerical scheme (e.g. of third or greater order): Discontinous Galerkin, Residual Distribution scheme, Spectral method, ...

This work extend interpolation error estimates to higher order numerical solution representation. We have examined the case of third-order accuracy. The first step is to reduce the tri-linear form given by the third order error term into a quadratic form based on the third order derivative. From this local error model, the optimal mesh is exhibited thanks to the continuous mesh framework.

3.15. Visualisation et modification des maillages courbes d'ordre élevé

Participants: Julien Castelneau, Adrien Loseille [correspondant], Loïc Maréchal.

Dans le cadre du projet ILab, des nouveaux algorithmes de visualisation et de modifications interactives des maillages courbes et hybrides ont été développés. En effet, une des principales difficultés dans la génération de maillages courbes reste la visualisation. Il est également nécessaire de disposer d'algorithmes de corrections interactifs car les maillages de surfaces initiaux (de degré 2) sont pour la plupart faux.

3.16. A changing-topology ALE numerical scheme

Participants: Frédéric Alauzet [correspondant], Nicolas Baral.

The main difficulty arising in numerical simulations with moving geometries is to handle the displacement of the domain boundaries, *i.e.,* the moving bodies. Only vertices displacement is not sufficient to achieve complex movement such as shear. We proved that the use of edge swapping allows us to achieve such complex displacement. We therefore developed an ALE formulation of this topological mesh modification to preserve the solver accuracy and convergence order. The goal is to extend to 3D the previous work done in 2D.

3.17. 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.

3.18. Maillages hexaédriques et calcul parallèle

Participant: Loïc Maréchal [correspondant].

Développement d'un remailleur de surfaces par la méthode octree. Celui-ci permet de passer d'une surface triangulée à problèmes (intersections de triangles, non-conformités, trous, etc.) à un maillage valide au sens des éléments finis.

Nouvelle version de la librairie d'aide au calcul sur GPU, GMLIB2, permettant de porter des codes travaillant sur des maillages de manière bien plus simple et efficace que la précédente. Des accélérations de l'ordre de 30, par rapport à un CPU en séquentiel, ont été obtenus avec le solveur Wolf et le mailleur Hexotic sur une carte Quadro 6000.

De nombreux développements sur le mailleur hexaédrique Hexotic ont été réalisés suite aux demandes de nombreux acheteurs industriels potentiels.

MATHRISK Team

5. New Results

5.1. Dynamic risk measures and BSDEs with jumps

The standard approach of mathematical quantification of financial risk in terms of Value at Risk has serious deficiencies. This has motivated a systematic analysis of risk measures which satisfy some minimal requirements of coherence and consistency. The theory of risk measures has been first developed in [54] in the coherent case and then extended in various directions (convex, dynamic, law-invariant) (see e.g. [70], [68], [93], [69]). We are extending this theory, in particular in the case of markets with possible random jumps and model ambiguity, and investigate various types of optimization problems involving risk measures.

Mathematical techniques for the treatment of such problems are based on non linear expectations, backward stochastic differential equations (BSDEs), stochastic control, stochastic differential games.

In the Brownian case, links between dynamic risk measures and Backward Stochastic Differential Equations (BSDEs) have been established (see, among others, [57]). A. Sulem and M.-C. Quenez are exploring these links in the case of stochastic processes with jumps. To this purpose, we have recently extended some comparison theorems for BSDEs with jumps given in [90], and provided a representation theorem of convex dynamic risk measures induced by BSDEs with jumps (see [44]). Optimization of dynamic risk measures leads to stochastic differential games or to optimal control problems for coupled systems of forward-backward stochastic differential equations (FBSDEs). They can be studied by stochastic maximum principles [100] or by transforming them into controlled Backward Stochastic Partial Differential Equations (BSPDEs). We address these questions in collaboration with B. Øksendal (Oslo university) and T. Zhang (Manchester University).

The numerical study of (F)BSDEs with jumps is especially demanding in high dimensions and collaboration has started on these issues with J. Lelong (ENSIMAG) and C. Labart (Université de Savoie).

5.2. Stochastic Differential Games

In many situations, controls are chosen by several agents who interact in various ways. To handle such cases one may use the theory of SDGs. This applies to model uncertainty problems, which can be regarded as a zero-sum game between the agent and the "market" and risk minimization, with risk represented via dynamic risk measures. More general non-zero sum games, involving several players, possibly with asymmetric information or delay will be studied.

An interesting new application of the theory of stochastic differential games, is the issue of *Public Private Partnership* which is a mechanism for a community to outsource the construction of public equipment. The community agrees to pay a rent to the contractor in order to cover the depreciation of the equipment, the maintenance costs and the financial costs. We want to model such partnerships and to compute and compare Nash equilibria and Stackelberg equilibria when the community is the leader. We would also like to investigate whether the community aversion to debt may lead it to enter such a partnership even this is more costly than constructing and managing the equipment by itself.

5.3. Optimal control of Stochastic Partial Differential equations (SPDEs)

SPDEs appear in the modeling of a number of situations: for example, in dynamic pollution models, in financial models involving interest rate derivatives, in systemic risk modeling. The research issues include optimal control of SPDEs and nonlinear filtering theory, stochastic control of forward-backward systems of SPDEs with imperfect and/or asymmetric information, optimal stochastic control of mean-field systems of SPDEs. We have started to study singular control of SPDEs. We plan to give a method for solving optimal control problems for general, possibly non-Markovian systems of FBSDEs by means of BSPDEs with jumps and associated comparison theorems.

5.4. Optimal stopping

Our research on optimal stopping problems covers the analysis of free boundaries in optimal stopping problems for multidimensional stochastic processes with jumps (Thesis of A. Bouselmi, supervised by D. Lamberton). Numerical issues are also be investigated (Monte Carlo methods, quantization methods, methods based on Malliavin calculus). Even in diffusion models, a realistic dividend modeling introduces jumps in the dynamics : at the dividend dates the spot value of the stock undergoes a jump equal to minus the dividend amount. We plan to take into account this feature in optimal stopping problems (Thesis of M. Jeunesse, supervised by B. Jourdain).

The pricing of American options with irregular payoff such as, for instance, binary options, leads to challenging mathematical problems. Some theoretical properties of optimal stopping problems with irregular payoffs have already been obtained. We now plan to focus on the Markovian case by using viscosity solutions and numerical analysis techniques.

In [45], we study optimal stopping problems for (non necessarily) convex dynamic risk measures induced by BSDEs with jumps and establish their connections with *Reflected* BSDEs with jumps. Such problems are related to optimal stopping for non linear expectations, which has been recently studied by [58] in the convex case only. We also address the case of model ambiguity and its relation with mixed control/optimal stopping problems.

5.5. Analysis of stochastic processes with jumps

The use of stochastic processes with jumps in financial modeling has been constantly increasing in the recent years. Simulation of these processes raises specific difficulties. A PhD thesis (V. Rabiet, adviser: V. Bally) has started on regularity properties of the law of multi-dimensional processes with jumps and on sensitivity analysis of derivative products with singular payoffs in such models.

5.6. Monte-Carlo methods

5.6.1. Adaptive variance reduction methods.

Stochastic algorithms [52], [53], [80], [63], [81], [27] or, more recently, direct stochastic optimization [73] proved to be a promising path to automatic variance reduction methods. Direct stochastic optimization techniques are easier to use in practice, avoiding completely any manual tuning needed for stochastic algorithms. This method is well understood (see [73]) only in the Gaussian case and for regular functions. We plan to extend the algorithms and prove rigorous results in non Gaussian cases with financial applications in view for jumps models (see [77], [76], [75]).

5.6.2. Monte-Carlo methods for calibration.

The interest for models combining local and stochastic volatility has been growing recently. Indeed, the local volatility model is not rich enough to efficiently deal with complex derivatives. A popular model is the so called Heston model, in which the volatility process solves a square-root stochastic differential equation (just as in the Cox-Ingersoll-Ross model for interest rate modeling). The thesis of L. Abbas-Turki [12](advisers: D. Lamberton and B. Lapeyre) concentrates on the multi-dimensional Heston model. For these models, numerical aspects are very demanding and we plan to use Monte-Carlo methods using advanced parallel devices (GPU clusters,...) both for price computations and calibration procedures. The thesis of Abbas-Turki is supported by the *Pôle de Compétitivité Finance Innovation* within the consortium *CrediNext*.

5.7. Systemic Risk

We extend the model in [51] in two major ways: First, study the optimal intervention strategy by a lender of last resort that would minimize the size of contagion under budget constraints. Second, allow our model not to be constrained to a single type of financial distress and model jointly insolvency and illiquidity. The interplay of these two mechanisms yields a more potent type of contagion than just the mechanical balance-sheet insolvency type of contagion [85], [92]. In [35], we have started to tackle these issues. This study can be enriched in many different manners.

Benjamin Jourdain and Agnès Sulem have organized a CEA-EDF-Inria school (70 participants) on the issues of Systemic risk and quantitative risk management in October 15-17 2012. (http://bit.ly/finance_inria). A special issue on "Systemic Risk" of the journal *Statistics and Risk Modeling* with B. Jourdain and A. Sulem as guest editors will be published in 2013.

MICMAC Project-Team

5. New Results

5.1. Electronic structure calculations

Participants: Eric Cancès, Ismaila Dabo, Virginie Ehrlacher, David Gontier, Salma Lahbabi, 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, V. Ehrlacher, S. Lahbabi and G. Stoltz have addressed issues related to the modeling and simulation of defects in periodic crystals.

Computing the energies of local defects in crystals is a major issue in quantum chemistry, materials science and nano-electronics. In collaboration with M. Lewin (CNRS, Cergy), E. Cancès and A. Deleurence have proposed in 2008 a new model for describing the electronic structure of a crystal in the presence of a local defect. This model is based on formal analogies between the Fermi sea of a perturbed crystal and the Dirac sea in Quantum Electrodynamics (QED) in the presence of an external electrostatic field. The justification of this model is obtained using a thermodynamic limit of Kohn-Sham type models. In [24], E. Cancès and G. Stoltz have studied the time evolution of defects within this model, in the context of linear response, which allowed them to give a rigorous meaning to the Adler-Wiser formula for the frequency-dependent dielectric permittivity of crystals. In collaboration with M. Lewin, E. Cancès and S. Lahbabi have introduced in [54] a functional setting for mean-field electronic structure models of Hartree-Fock or Kohn-Sham types for disordered quantum systems, and used these tools to study the reduced Hartree-Fock model for a disordered crystal where the nuclei are classical particles whose positions and charges are random.

On the numerical side, E. Cancès has worked with Y. Maday and R. Chakir (University Paris 6) on the numerical analysis of the electronic structure models. In [22], they have obtained optimal *a priori* error bounds for the the planewave approximation of the Thomas-Fermi-von Weizsäcker and the Kohn-Sham LDA models. Together with Y. Maday, E. Cancès and V. Ehrlacher have analyzed the computation of eigenvalues in spectral gaps of locally perturbed periodic Schrödinger operators [23]. In [53], they have introduced a general theoretical framework to analyze non-consistent approximations of the discrete eigenmodes of a self-adjoint operator, focusing in particular on the discrete eigenvalues laying in spectral gaps. Applying this analysis to the supercell method for perturbed periodic Schrödinger operators, they derive optimal convergence rates for the planewave discretization method, taking numerical integration errors into account. These results, along with earlier work on greedy algorithms for nonlinear convex problems and the study of local defects in the Thomas-Fermi-von Weiszacker theory, are collected in [7].

In the work [38], Claude Le Bris, in collaboration with Pierre Rouchon (Ecole des Mines de Paris), has introduced 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: Matthew Dobson, Claude Le Bris, Frédéric Legoll, Tony Lelièvre, Francis Nier, Grigorios Pavliotis, Mathias Rousset, Gabriel Stoltz.

The extremely broad field of molecular dynamics is a domain where the MICMAC project-team, originally more involved in the quantum chemistry side, has invested a lot of efforts in the recent years. Molecular dynamics may also be termed computational statistical physics since the main aim is to numerically estimate average properties of materials as given by the laws of statistical physics. The project-team studies both deterministic and probabilistic techniques used in the field. One of the main difficulty is related to the metastable features of the generated trajectories: the system remains trapped over very long times in metastable states, which means that very long trajectories need to be generated in order to obtain macroscopically relevant quantities. This is related to the fact that the timescale at the microscopic level is much smaller than the timescale at the macroscopic level. In [66], we propose a summary of the mathematical approaches to quantify metastability, and which appear to be useful to analyze the numerical methods used in molecular dynamics.

5.2.1. Free Energy calculations

For large molecular systems, the information of the whole configuration space may be summarized in a few coordinates of interest, called reaction coordinates. An important problem in chemistry or biology is to compute the effective energy felt by those reaction coordinates, called free energy.

In the article [42], Tony Lelièvre, Mathias Rousset and Gabriel Stoltz study the application of constrained Langevin dynamics to the computation of free energy differences, by thermodynamic integration techniques and fluctuation relation (à la Jarzynski).

One interest of free energy computation techniques is that they appear to be useful in other fields, like in computational statistics where multimodal measures are also frequently encountered, so that standard Markov Chain Monte Carlo appraoches also suffer from metastability.

For example, in [25], Nicolas Chopin (CREST, ENSAE), T. Lelièvre and G. Stoltz explore the application of the Adaptive Biasing Force method to Bayesian inference. This sampling method belongs to the general class of adaptive importance sampling strategies which use the free energy along a chosen reaction coordinate as a bias. Such algorithms are very helpful to enhance the sampling properties of Markov Chain Monte Carlo algorithms, when the dynamic is metastable.

In [58], G. Fort (Telecom Paris), B. Jourdain (CERMICS), E. Kuhn (INRA), T. Lelièvre and G. Stoltz have considered the Wang-Landau algorithm. The authorshave proved that the Wang-Landau algorithm converges with an associated central limit theorem, and have provided an analysis of the efficiency of the algorithm in a metastable situation.

5.2.2. Convergence to equilibrium

An important question for the analysis of sampling techniques is the rate of convergence to equilibrium for stochastic trajectories.

In [65], F. Nier, T. Lelièvre and G. Pavliotis study the interest of using non-reversible stochastic dynamics to enhance the rate of convergence to equilibrium, compared to reversible dynamics. A well posed optimization problem is obtained and solved in the case of a linear drift for the overdamped Langevin dynamics.

5.2.3. Metropolis Hastings algorithms

A classical sampling tool used in molecular dynamics and in computational statistics is the Metropolis-Hastings algorithm. There has been a lot of work (see G. Roberts et al.) to study how the variance of the proposal should scale with the dimension of the problem, in order to optimize the sampling procedure. Most of these works assume that (i) the target probability is the product of n one dimensional laws and that (ii) the Markov chain starts at equilibrium.

In the two works [60], [59], T. Lelièvre and his co-authors have generalized these results when the initial distribution is not the target probability. The diffusive limit in the latte case is solution to a stochastic differential equation nonlinear in the sense of McKean. They have discussed 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 $1/3$) first suggested in the works of G. Roberts et al., at least for the Random Walk Metropolis algorithm.

5.2.4. Thermodynamic limit

The quasicontinuum method is an approach to couple an atomistic model with a coarse-grained approximation in order to compute the states of a crystalline lattice at a reduced computational cost compared to a full atomistic simulation.

In that framework, the team has addressed questions related to the *finite temperature* modeling of atomistic systems and derivation of coarse-grained descriptions, such as canonical averages of observables depending only on a few variables. In the one-dimensional setting, an efficient strategy that bypasses the simulation of the whole system had been proposed in 2010. We refer to [47] for a recent review. In collaboration with X. Blanc (Université Pierre et Marie Curie), F. Legoll has extended this strategy to the so-called membrane setting in [16].

When the temperature is small, a perturbation approach can be used to compute the canonical averages of these observables depending only on a few variables, at first order with respect to temperature. In collaboration with E. Tadmor, W. K. Kim, L. Dupuy and R. Miller, F. Legoll has analyzed such an approach in [46]. The numerical tests reported there show the efficiency of the approach, as long as the temperature is indeed small.

5.2.5. Sampling trajectories

There exist a lot of methods to sample efficiently Boltzmann-Gibbs distributions. The situation is much more intricated as far as the sampling of trajectories (and especially metastable trajectories) is concerned.

Following a numerical observation in a previous work on the sampling of reactive trajectories by a multilevel splitting algorithm, F. Cérou (Inria Rennes), A. Guyader (Inria Rennes), T. Lelièvre and F. Malrieu (Université de Rennes) study theoretically in [56] the distribution of the lengths of these trajectories, using large deviation techniques.

In [37], C. Le Bris and T. Lelièvre together with M. Luskin and D. Perez from Los Alamos National Laboratoy provide a mathematical analysis of the parallel replica algorithm, which has been proposed by A. Voter in 1997 to simulate very efficiently metastable trajectories. This work opens a lot of perspectives, by using a generic tool (the quasi stationary distribution) to make a link between a continuous state space dynamics (Langevin dynamics) and a discrete state space dynamics (kinetic Monte Carlo models).

In a work in progress, T. Lelièvre and F. Nier have studied the quasi-stationnary distribution in relation for an overdamped Langevin process in a bounded domain. In the small temperature limit and by making the connection with boundary Witten Laplacians, they are able to compute accurately the spatial exit law along the boundary and non perturbative accurate formulas when the potential is changed inside the domain.

5.2.6. Effective dynamics

For a given molecular system, and a given reaction coordinate $\xi : \mathbb{R}^n \mapsto \mathbb{R}$, the free energy completely describes the statistics of $\xi(X)$ when $X \in \mathbb{R}^n$ is distributed according to the Gibbs measure. On the other hand, obtaining a correct description of the dynamics along ξ is complicated.

F. Legoll and T. Lelièvre have introduced and analyzed some years ago a strategy to define a coarse-grained dynamics that approximates $\xi(X_t)$, when the state of the system X_t evolves according to the overdamped Langevin equation (which is ergodic for the Gibbs measure). We refer to [47] for a recent review. The aim was to get a coarse-grained description giving access to some *dynamical* quantities (and not only *equilibrium* quantities). Together with G. Samaey (KU Leuven), they have recently studied how to use this coarse-grained description, accurate when the time scale separation is asymptotically large, to somewhat precondition the dynamics of the actual system in cases when the time scale separation is not large. For that purpose, they have used the parareal framework, to iteratively correct the sequential coarse-grained trajectory by fine scale trajectories performed in parallel. The main difficulty is that the two models (the reference one and the coarse-grained one) do not act on the same variable: the reference model evolves all the variables, whereas the coarse-grained model only evolves the slow variables. As shown in [63] in a simplified context (that of singularly perturbed ODEs), the precise coupling between both models should be done carefully.

The above study is concerned with models with continuous state spaces. S. Lahbabi and F. Legoll have studied in [61] a related question in the framework of kinetic Monte Carlo models, where the state space is discrete. For some models involving some slow and some fast variables, the effective dynamics of the slow component has been identified, and a complete proof of convergence proposed.

5.2.7. Hamiltonian dynamics

Constant energy averages are often computed as long time limits of time averages along a typical trajectory of the Hamiltonian dynamics. One difficulty of such a computation is the presence of several time scales in the dynamics: the frequencies of some motions are very high (e.g. for the atomistic bond vibrations), while those of other motions are much smaller. This problem has been addressed in a two-fold manner.

Fast phenomena are often only relevant through their mean effect on the slow phenomena, and their precise description is not needed. Consequently, there is a need for time integration algorithms that take into account these fast phenomena only in an averaged way, and for which the time step is not restricted by the highest frequencies. In [29], M. Dobson, C. Le Bris, and F. Legoll have developed integrators for Hamiltonian systems with high frequencies. The integrators were derived using homogenization techniques applied to the Hamilton-Jacobi PDE associated to the Hamiltonian ODE. This work extends previous works of the team. The proposed algorithms can now handle the case when the (unique) fast frequency depends on the slow degrees of freedom, or when there are several fast constant frequencies.

Another track to simulate the system for longer times is to resort to parallel computations. An algorithm in that vein is the parareal in time algorithm. It is based on a decomposition of the time interval into subintervals, and on a predictor-corrector strategy, where the propagations over each subinterval for the corrector stage are concurrently performed on the processors. Using a symmetrization procedure and/or a (possibly also symmetric) projection step, C. Le Bris and F. Legoll, in collaboration with X. Dai and Y. Maday, have introduced several variants of the original plain parareal in time algorithm [28]. These variants, compatible with the geometric structure of the exact dynamics, are better adapted to the Hamiltonian context.

5.2.8. Nonequilibrium systems

The efficient simulation of molecular systems is known to be a much more complicated problem when the system is subjected to a non-conservative external forcing than when the system experiences conservative forces. Together with the sampling of metastable dynamics mentioned above, these are the two major research focus in molecular dynamics of the project-team.

Nonequilibrium molecular dynamics simulations can be used to compute the constitutive relation between the strain rate and stress tensor in complex fluids. This is fulfilled simulating molecular systems subject to a steady, non-zero macroscopic flow at a given temperature. Starting from a bath model, M. Dobson, F. Legoll, T. Lelièvre, and G. Stoltz have derived a Langevin-type dynamics for a heavy particle in a non-zero background flow [57]. The resulting dynamics, which is theoretically obtained when a *unique* large particle is considered, is numerically observed to also perform well when a *system* of many interacting particles within shear flow is considered.

Let us also mention that the article on the computation of the viscosity of fluids using steady state nonequilibrium dynamics with an external nongradient bulk forcing, in the framework of the PhD of Rémi Joubaud, has also been published [34]. In addition, the study by G. Stoltz and C. Bernardin on thermal transport in one-dimensional chains of oscillators whose kinetic and potential energy functions are the same, has been accepted and is now published [13].

5.3. Complex fluids

Participants: David Benoit, Sébastien Boyaval, Claude Le Bris, Tony Lelièvre.

In [41], Claude Le Bris and Tony Lelièvre review the state-of-the-art of numerical and mathematical results on micro-macro models for viscoelastic fluids.

Following previous works, in [32], Claude Le Bris and Tony Lelièvre together with Lingbing He analyze the longtime behaviour of nematic polymeric fluids (liquid crystals). The longtime asymptotic for such models is much richer than for flexible polymers, that were considered in a previous analysis. Indeed, for these models, periodic in time behaviours are observed.

In his PhD under the supervision of Claude Le Bris and Tony Lelièvre, David Benoît studies models of aging fluids developed at the ESPCI (Ecole supérieure de physique et de chimie industrielles) and designed to take into account phenomena such as shear thinning, aging and shear banding in falling sphere experiments. The work consists in studying on the one hand the mathematical well-posedness of some macroscopic models, see [51] and, on the other hand, in trying to understand the link between such macroscopic models and microscopic models which have been proposed to describe such fluids.

Related to the mathematical modelling of free-surface complex flows under gravity, a new reduced model for thin layers of a viscoelastic upper-convected Maxwell fluid was derived by S. Boyaval in collaboration with François Bouchut, and possibly discontinuous solutions were numerically simulated with a new finite-volume scheme of relaxation type that satisfies a discrete counterpart of the natural dissipation [20]. This work is being pursued for other models.

Finally, in [31], Alexandre Ern (CERMICS), Rémi Joubaud (CERMICS) and Tony Lelièvre analyze a model describing equilibrium binary electrolytes surrounded by charged solid walls. This work is done in collaboration with physicists from the group PECSA at Université Pierre et Marie Curie. Applications include the modelization of clays for the burying of nuclear waste.

5.4. Application of greedy algorithms

Participants: Sébastien Boyaval, Eric Cancès, Virginie Ehrlacher, Tony Lelièvre.

Greedy algorithms are used in many contexts for the approximation of high-dimensional functions: Proper Generalized Decomposition, Reduced Basis techniques, etc.

Various greedy algorithms for high-dimensional non-symmetric problems, and inherent theoretical and practical difficulties have been analyzed in [52]. Current research now aims at extending these techniques to the approximation of high-dimensional spectral problems. Prototypical applications include electronic structure calculations or the computation of buckling modes in mechanics.

In probabilistic methods for uncertainty quantification in mechanics, S. Boyaval has used a greedy algorithm to construct control variates for accelerating Monte-Carlo simuation in the cases where an expectation has to be computed many times [21]. The work is being applied to the uncertainty quantification in numerical models for hydraulic engineering.

Finally, in [55], Fabien Casenave (CERMICS), Alexandre Ern (CERMICS) and Tony Lelièvre study the influence of round-off errors on the evaluation of the a posteriori estimators in the reduced basis approach. In practice, the evaluation of the error estimator can become very sensitive to round-off errors. An explanation of this fact is proposed, as well as efficient remedies.

5.5. Mathematical Physics

Participant: Francis Nier.

In [10], A. Aftalion and F. Nier answer questions asked by J. Dalibard about the feasibility of artificial gauge potentials. This analysis provides the range of small parameters within which the linear adiabatic argument used by the physicists is certainly not destroyed by the non linear effects.

In [43], D. Le Peutrec, F. Nier and C. Viterbo give an accurate Arrhenius law for Witten Laplacian acting on p-forms. In the case of functions the exponentially small eigenvalues are given by exponentiated differences of enegy levels between local minima and saddle points (Arrhenius law). In the case of p-forms the association of critical points with index p and critical points with index p+1 or p-1, is more subtle and is provided by Barannikov's presentation of Morse theory.

In [11], Z. Ammari and F. Nier have proved the mean field dynamics of general bosonic systems in the presence of singular pair interaction potentials, including the important 3 dimensional Coulombic case. As compared with their previous works, they developed a slightly new strategy relying on measure transportation techniques and results presented by Ambrosio-Gigli-Savaré in their book "Gradient Flows: In Metric Spaces And In The Space Of Probability Measures" (2005).

5.6. Homogenization and related topics

Participants: Ronan Costaouec, Claude Le Bris, Frédéric Legoll, William Minvielle, Mathias Rousset, Florian Thomines.

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 in [17] 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. Promising results have been obtained on a simple case of a periodic system perturbed by a localized defect. Ongoing works consider other configurations, such as for instance an interface between two different crystalline phases.

A theme closely related to homogenization theory and on which several members of the project team have worked a lot in the past few years is the passage from discrete (atomistic) mechanics to continuum mechanics. In this direction, C. Le Bris, in collaboration with X. Blanc and P.-L. Lions, has established in [18] the rigorous continuum limit of the Newton equations of motion for some simple cases of one-dimensional atomistic system.

The project-team also 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.

An interesting case in that context is when the randomness comes as a *small* perturbation of the deterministic case. As previously shown by earlier works of the project-team, this situation can indeed be handled with a dedicated approach, which turns out to be far more efficient than the standard approach of stochastic homogenization. A final component of the work completed by Florian Thomines during his PhD thesis has concerned the application of Reduced Basis techniques to that specific context of weakly stochastic homogenization problems. In particular, the approach has been adapted in [39] to efficiently compute the terms of the expansion previously developed by A. Anantharaman and C. Le Bris to approximate a certain category of weakly random homogenization problems. It has been demonstrated that the reduced basis technique is very helpful in this particular context and indeed allows for a speed up of the computation. Another application of the same technique, for a slightly different category of models (still in the framework of weakly random homogenization problems) originally derived by X. Blanc, P.-L. Lions and C. Le Bris, has also been explored. The difficulty, there, is to compute the various corrector equations that parametrically depend on the macroscopic location of the microstructure and the particular realization of that microstructure. The problem is definitely amenable to reduced basis techniques, as demonstrated by some preliminary tests, but definite conclusions on the general validity of the approach are yet to be obtained.

The team has also proceeded to address, from a numerical viewpoint, the case when the randomness is not small. In that case, 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. For a given truncation of $\mathbb{R}^d$, the team has shown in [14] that the variance of this matrix can be reduced using the technique of antithetic variables. F. Legoll and W. Minvielle are currently extending this technique to nonlinear, convex homogenization problems.

In addition, C. Le Bris, F. Legoll, W. Minvielle and M. Rousset are currently investigating the possibility to use other variance reduction approaches, such as control variate techniques. A promising idea is to use the weakly stochastic model previously introduced by A. Anantharaman and C. Le Bris (in which a periodic model is perturbed by a *rare* stochastic perturbation) to build a control variate model. The preliminary results that have already been obtained are very encouraging.

Another contribution in stochastic homogenization is the following. C. Le Bris, in collaboration with X. Blanc and P.-L. Lions, has recently introduced a variant of the classical random homogenization. For that variant, as often in random homogenization, the homogenized matrix is again defined from a so-called corrector function, which is the solution to a problem set on the entire space. F. Legoll and F. Thomines have described and proved the almost sure convergence of an approximation strategy based on truncated versions of the corrector problem in [64]. F. Legoll and F. Thomines have also established, in the one-dimensional case, a convergence result on the residual process, defined as the difference between the solution to the highly oscillatory problem and the solution to the homogenized problem.

From a numerical perspective, the Multiscale Finite Element Method 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 extension of this strategy to the stochastic case, when the tensor describing the properties of the material is the sum of a periodic term and a small random term, has been studied by C. Le Bris, F. Legoll and F. Thomines [36]. A method with a much smaller computational cost than the original MsFEM in the stochastic setting has been proposed. Provided the stochastic perturbation is indeed small, the proposed method is as accurate as the original one. The work [36] also provides a complete analysis of the approach, extending that available for the deterministic setting. Such analysis often rely on the rate of convergence of the two scale expansion (in the sense of homogenization theory) of the solution to the highly oscillatory elliptic partial differential equation. Such a result is classic for periodic homogenization. In generic stochastic homogenization, the rate can be arbitrary small, depending on the rate with which the correlations of the random coefficient vanish. C. Le Bris, F. Legoll and F. Thomines have established in [40] such a result for *weakly stochastic homogenization*, using asymptotic properties of the Green function of the elliptic operator $Lu = -\text{div}\ (A\nabla u)$ (where A is a periodic, coercive and bounded matrix), established by F. Legoll in collaboration with X. Blanc and A. Anantharaman [15].

Still in the framework of the Multiscale Finite Element approach, F. Thomines has further investigated, in collaboration with Y. Efendiev and J. Galvis (Texas A&M University), the use of Reduced Basis methods. They have considered an extension of the MsFEM approach, well suited to the high contrast case, i.e. the case when the ratio between the maximum and the minimum values of the heterogeneous coefficient is large. The main idea of this extension is to complement the standard MsFEM basis functions with the eigenfunctions (associated to the first small eigenvalues) of a local eigenvalue problem. In [30], Y. Efendiev, J. Galvis and F. Thomines have considered the case when the problem depends on an additional parameter, and have shown how to use the Reduced Basis approach to more efficiently compute the eigenfunctions mentioned above.

Even in simple deterministic cases, there is actually still room for improvement in many different directions for the MsFEM approach. In collaboration with A. Lozinski (University of Toulouse and now at the University of Besançon) who visited the team-project repeatedly during the year, 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. The approach has been analyzed (combining classical arguments from homogenization theory and finite element theory) and tested on simple, but highly convincing cases [35]. In particular, an elliptic problem set on a domain with a huge number of perforations has been considered in [62]. 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, two-week visitor in the project-team in the Spring of 2012), consists in the study of multiscale advection-diffusion problems. Such problems are possibly advection dominated and a stabilization procedure is therefore required. How stabilization interferes with the multiscale character of the equation is an unsolved mathematical question worth considering for numerical purposes.

Still another question related to homogenization theory that is investigated in the group is the following. Consider an elliptic equation, say in divergence form, with a highly oscillatory matrix coefficient. Is it possible to approximate the boundary value problem for different right hand sides using a similar problem with a *constant* matrix coefficient? How can this "best" constant matrix approximating the oscillatory problem be constructed in an efficient manner? How is this matrix related to the homogenized matrix, in the limit of infinitely rapidly oscillatory coefficients? Current work is directed towards solving such issues.

5.7. Asymptotic variance reduction

Participant: Mathias Rousset.

Recently, M. Rousset has initiated a research topic on variance reduction techniques (called "asymptotic") for the simulation of stochastic models of particles. The point is to use a macroscopic (or model reduced) equation as a control variate; or in other words, to use the information of a macroscopic description to decrease the statistical error of the simulated microscopic evolution.

A first step in this program has been achieved for a microscopic model describing the individual motion of bacteriae with a Markovian velocity-jump process. The macroscopic equation is an advection-diffusion equation called the chemotaxis equation. In [44], the pobabilistic derivation of the chemotaxis equation from the individual motion of bacteriae have been carried out in a rigorous way. In [45], a numerical method simulating the individual evolution of bacteriae with "asymptotic" variance reduction have been proposed.

5.8. Computational materials spectroscopy in electrochemistry and optoelectronics

Participant: Ismaila Dabo.

Many advances in the understanding and design of nanomaterials have been enabled by spectroscopic techniques of increasing spatial and temporal resolution. In electrochemistry and optoelectronics, spectroscopy provides insight into the chain of processes involved in harnessing, storing, and delivering energy.

In support to experimental techniques, much progress has been achieved in simulating spectroscopic phenomena to shed light into energy conversion at the molecular scale. Such understanding is critical to the molecular design of a range of electrical devices, including but not limited to fuel cells, batteries, dye-sensitized solar cells, and optoelectronic devices.

The work of I. Dabo is dedicated to the development of quantum and semiclassical methods to simulate spectroscopies of electrochemical and optoelectronic materials. The achieved level of efficiency and accuracy fosters dialogue between experiment and theory for interpreting complex spectroscopic data. This year, these novel methods have been applied to simulate spectroscopic phenomena spanning the infrared to the visible and ultraviolet ranges.

The first application pertains to the infrared sum-frequency-generation (SFG) spectroscopy of adsorption mechanisms at the origin of the tolerance of fuel-cell catalytic electrodes to chemical poisoning. The study explains the critical influence of the electrode voltage in analyzing surface spectroscopy experiments (work done in collaboration with EPFL). [12], [26], [19]

The second application aims at understanding the sensitizing properties of organometallic dyes in dye-sensitized solar cells by simulating optical photoluminescence (PL) spectra, thereby elucidating the role of electron localization and ligand functionalization on the phosphorescence of organometallic complexes (work done in collaboration with the University of Minnesota). [33]

The third application is focused on the ultraviolet photoelectron spectroscopy (UPS) of photoactive nanomaterials of relevance to the design of organic photovoltaic junctions and photoelectrodes (work done in collaboration with the Italian Institute of Nanoscience, Seoul National University, and Xiamen University). [27]

Future challenges and opportunities are related to the time-dependent simulation of transient and cyclic spectra. These developments, which will be part of the widely used Quantum-ESPRESSO distribution (http://www.quantum-espresso.org), would pave the way for comprehensive studies of kinetic processes in tandem with time-resolved spectroscopic experiments.

SIERRA Project-Team

6. New Results

6.1. A Stochastic Gradient Method with an Exponential Convergence Rate for Strongly-Convex Optimization with Finite Training Sets

Participants: Francis Bach, Mark Schmidt, Nicolas Le Roux [correspondant].

In [21], we propose a new stochastic gradient method for optimizing the sum of a finite set of smooth functions, where the sum is strongly convex. While standard stochastic gradient methods converge at sublinear rates for this problem, the proposed method incorporates a memory of previous gradient values in order to achieve a linear convergence rate. In a machine learning context, numerical experiments indicate that the new algorithm can dramatically outperform standard algorithms, both in terms of optimizing the training objective and reducing the testing objective quickly.

6.2. Convex Relaxation for Combinatorial Penalties

Participants: Francis Bach, Guillaume Obozinski [correspondant].

In [15], we propose an unifying view of several recently proposed structured sparsity-inducing norms. We consider the situation of a model simultaneously (a) penalized by a set- function de ned on the support of the unknown parameter vector which represents prior knowledge on supports, and (b) regularized in Lp-norm. We show that the natural combinatorial optimization problems obtained may be relaxed into convex optimization problems and introduce a notion, the lower combinatorial envelope of a set-function, that characterizes the tightness of our relaxations. We moreover establish links with norms based on latent representations including the latent group Lasso and block-coding, and with norms obtained from submodular functions.

6.3. Kernel change-point detection

Participant: Sylvain Arlot [correspondant].

In [16], we tackle the change-point problem with data belonging to a general set. We propose a penalty for choosing the number of change-points in the kernel-based method of Harchaoui and Cappé (2007). This penalty generalizes the one proposed for one dimensional signals by Lebarbier (2005). We prove it satisfies a non-asymptotic oracle inequality by showing a new concentration result in Hilbert spaces. Experiments on synthetic and real data illustrate the accuracy of our method, showing it can detect changes in the whole distribution of data, even when the mean and variance are constant. Our algorithm can also deal with data of complex nature, such as the GIST descriptors which are commonly used for video temporal segmentation.

Collaboration with Alain Celisse (University Lille 1; Inria Lille, MODAL team) and Zaïd Harchaoui (Inria Grenoble, LEAR team).

6.4. On the Equivalence between Herding and Conditional Gradient Algorithms

Participants: Francis Bach [correspondant], Simon Lacoste-Julien, Guillaume Obozinski.

In [5], we show that the herding procedure of Welling (2009) takes exactly the form of a standard convex optimization algorithm–namely a conditional gradient algorithm minimizing a quadratic moment discrepancy. This link enables us to invoke convergence results from convex optimization and to consider faster alternatives for the task of approximating integrals in a reproducing kernel Hilbert space. We study the behavior of the different variants through numerical simulations. The experiments indicate that while we can improve over herding on the task of approximating integrals, the original herding algorithm tends to approach more often the maximum entropy distribution, shedding more light on the learning bias behind herding.

6.5. V-fold cross-validation and V-fold penalization in least-squares density estimation

Participant: Sylvain Arlot [correspondant].

In [22], we study V-fold cross-validation for model selection in least-squares density estimation. The goal is to provide theoretical grounds for choosing V in order to minimize the least-squares risk of the selected estimator. We first prove a non asymptotic oracle inequality for V-fold cross-validation and its bias-corrected version (V-fold penalization), with an upper bound decreasing as a function of V. In particular, this result implies V-fold penalization is asymptotically optimal. Then, we compute the variance of V-fold cross-validation and related criteria, as well as the variance of key quantities for model selection performances. We show these variances depend on V like $1 + 1/(V - 1)$ (at least in some particular cases), suggesting the performances increase much from $V = 2$ to $V = 5$ or 10, and then is almost constant. Overall, this explains the common advice to take $V = 10$—at least in our setting and when the computational power is limited—, as confirmed by some simulation experiments.

Collaboration with Matthieu Lerasle (CNRS, University Nice Sophia Antipolis).

6.6. Machine learning for Neuro-imaging

Participants: Fabian Pedregosa [correspondant], Francis Bach, Guillaume Obozinski.

In the course of the year 2011-2012 two articles where submitted and accepted in international workshops. The first published article, **Improved brain pattern recovery through ranking approaches** ([12]) was presented at the 2nd International Workshop on Pattern Recognition in NeuroImaging in London, July 2012 and proposes a new approach for the problem of estimating the coefficients of a generalized linear model with monotonicity constraint. For this, we explore the use of ranking techniques, which are popular in the context of information retrieval but novel for medical imaging applications.

The second published article, **Learning to rank from medical imaging data** ([11]) uses the same techniques as the previous article to solve a more fundamental problem, that is, to predict a quantitative (and potentially non-linear) variable from a set of noisy measurements. We show on simulations and two fMRI datasets that this approach is able to predict the correct ordering on pairs of images, yielding higher prediction accuracy than standard regression and multiclass classification techniques.

Collaboration with the Parietal project-team (A. Gramfort, B. Thirion, G. Varoquaux)

6.7. SiGMa: Simple Greedy Matching for Aligning Large Knowledge Bases

Participant: Simon Lacoste-Julien [correspondant].

The Internet has enabled the creation of a growing number of large-scale knowledge bases in a variety of domains containing complementary information. Tools for automatically aligning these knowledge bases would make it possible to unify many sources of structured knowledge and answer complex queries. However, the efficient alignment of large-scale knowledge bases still poses a considerable challenge. In [20], we present Simple Greedy Matching (SiGMa), a simple algorithm for aligning knowledge bases with millions of entities and facts. SiGMa is an iterative propagation algorithm which leverages both the structural information from the relationship graph as well as flexible similarity measures between entity properties in a greedy local search, thus making it scalable. Despite its greedy nature, our experiments indicate that SiGMa can efficiently match some of the world's largest knowledge bases with high precision. We provide additional experiments on benchmark datasets which demonstrate that SiGMa can outperform state-of-the-art approaches both in accuracy and efficiency.

Collaboration with Konstantina Palla, Alex Davies, Zoubin Ghahramani (Machine Learning Group, Department of Engineering, University of Cambridge); Gjergji Kasneci (Max Planck Institut fur Informatik); Thore Graepel (Microsoft Research Cambridge).

6.8. Block-Coordinate Frank-Wolfe Optimization for Structural SVMs

Participants: Simon Lacoste-Julien [correspondant], Mark Schmidt.

In [19], we propose a randomized block-coordinate variant of the classic Frank-Wolfe algorithm for convex optimization with block-separable constraints. Despite its lower iteration cost, we show that it achieves the same convergence rate in duality gap as the full Frank-Wolfe algorithm. We also show that, when applied to the dual structural support vector machine (SVM) objective, this yields an online algorithm that has the same low iteration complexity as primal stochastic subgradient methods. However, unlike stochastic subgradient methods, the stochastic Frank-Wolfe algorithm allows us to compute the optimal step-size and yields a computable duality gap guarantee. Our experiments indicate that this simple algorithm outperforms competing structural SVM solvers.

Collaboration with Martin Jaggi (Centre de Mathématiques Appliquées, Ecole Polytechnique); Patrick Pletscher (Machine Learning Laboratory, ETH Zurich).

6.9. A convex relaxation for weakly supervised classifiers

Participants: Armand Joulin [correspondant], Francis Bach.

In [8], we introduce a general multi-class approach to weakly supervised classification. Inferring the labels and learning the parameters of the model is usually done jointly through a block-coordinate descent algorithm such as expectation-maximization (EM), which may lead to local minima. To avoid this problem, we propose a cost function based on a convex relaxation of the soft-max loss. We then propose an algorithm specifically designed to efficiently solve the corresponding semidefinite program (SDP). Empirically, our method compares favorably to standard ones on different datasets for multiple instance learning and semi-supervised learning, as well as on clustering tasks.

6.10. Multi-Class Cosegmentation

Participants: Armand Joulin [correspondant], Francis Bach.

Bottom-up, fully unsupervised segmentation remains a daunting challenge for computer vision. In the cosegmentation context, on the other hand, the availability of multiple images assumed to contain instances of the same object classes provides a weak form of supervision that can be exploited by discriminative approaches. Unfortunately, most existing algorithms are limited to a very small number of images and/or object classes (typically two of each). In [9], we propose a novel energy-minimization approach to cosegmentation that can handle multiple classes and a significantly larger number of images. The proposed cost function combines spectral- and discriminative-clustering terms, and it admits a probabilistic interpretation. It is optimized using an efficient EM method, initialized using a convex quadratic approximation of the energy. Comparative experiments show that the proposed approach matches or improves the state of the art on several standard datasets.

Collaboration with the Willow project-team (J. Ponce).

6.11. A latent factor model for highly multi-relational data

Participants: Nicolas Le Roux, Guillaume Obozinski [correspondant].

Many data such as social networks, movie preferences or knowledge bases are multi-relational, in that they describe multiple relations between entities. While there is a large body of work focused on modeling these data, modeling these multiple types of relations jointly remains challenging. Further, existing approaches tend to breakdown when the number of these types grows. In [7], we propose a method for modeling large multi relational datasets, with possibly thousands of relations. Our model is based on a bilinear structure, which captures various orders of interaction of the data, and also shares sparse latent factors across different relations. We illustrate the performance of our approach on standard tensor-factorization datasets where we attain, or outperform, state-of-the-art results. Finally, a NLP application demonstrates our scalability and the ability of our model to learn efficient and semantically meaningful verb representations.

Collaboration with R. Jenatton (CMAP, Ecole Polytechnique) and Antoine Bordes (CNRS, Université de Technologie de Compiégne).

6.12. Semi-supervised NMF with time-frequency annotations for single-channel source separation

Participants: Francis Bach, Augustin Lefèvre [correspondant].

In [10], we formulate a novel extension of nonnegative matrix factorization (NMF) to take into account partial information on source-specific activity in the spectrogram. Results on single-channel source separation show that time-frequency annotations allow to disambiguate the source separation problem, and learned annotations open the way for a completely unsupervised learning procedure for source separation with no human intervention.

Collaboration with C. Févotte (Laboratoire traitement et communication de l'information (LTCI), CNRS: UMR5141 - Institut Télécom - Télécom ParisTech).

BANG Project-Team

6. New Results

6.1. Proliferation dynamics and its control

6.1.1. Cell division dynamics in structured cell populations

Participants: José Luís Avila Alonso [DISCO project-team, Inria Saclay IdF], Annabelle Ballesta, Frédérique Billy, Frédéric Bonnans [Commands project-team, Inria Saclay IdF], Catherine Bonnet [DISCO project-team, Inria Saclay IdF], Jean Clairambault, Luna Dimitrio, Marie Doumic-Jauffret, Xavier Dupuis [Commands project-team], Olivier Fercoq [MaxPlus project-team, Inria Saclay IdF], Stéphane Gaubert [MaxPlus project-team, Inria Saclay IdF], Germain Gillet [IBCP, Université Cl. Bernard Lyon 1], Philippe Gonzalo [IBCP, Université Cl. Bernard Lyon 1], Pierre Hirsch [INSERM Paris (Team18 of UMR 872) Cordeliers Research Centre and St. Antoine Hospital, Paris], Thomas Lepoutre [now in DRACULA project-team, Inria Rhône-Alpes, Lyon], Jonathan Lopez [IBCP, Université Cl. Bernard Lyon 1], Pierre Magal [University Bordeaux II], Anna Marciniak-Czochra [Institute of Applied Mathematics, Universität Heidelberg], Jean-Pierre Marie [INSERM Paris (Team18 of UMR 872) Cordeliers Research Centre and St. Antoine Hospital, Paris], Roberto Natalini [IAC-CNR, Università Sapienza, Rome], Silviu Niculescu [DISCO project-team, Inria Saclay IdF], Hitay Özbay [Bilkent University, Ankara, Turkey], Benoît Perthame, Ruoping Tang [INSERM Paris (Team18 of UMR 872) Cordeliers Research Centre and St. Antoine Hospital, Paris], Vitaly Volpert [CNRS Lyon, UMR5208, Camille Jordan Institute, Lyon], Jorge Zubelli [IMPA, Rio de Janeiro].

1. *Transition kernels in a McKendrick model of the cell division cycle.* This theme has continued to be developed with identification of model parameters by FUCCI imaging in collaboration with G. van der Horst's team in Amsterdam and with F. Delaunay's team in Nice, within the C5Sys European network, coordinated by F. Lévi (Villejuif) [10], [11], [12], [39], [42], [43]. Main young researchers on this theme, F. Billy has concluded her 2-year Inria postdoc at Bang, leaving for an industrial company in November 2012, and O. Fercoq (team MaxPlus, Saclay) has defended his PhD thesis at École Polytechnique in September 2012, only to leave for a postdoc position dedicated to optimisation theory in Edinburgh.
2. *Modelling haematopoiesis with applications to AML.* This theme has been active through a collaboration with Inria teams Commands (F. Bonnans, X. Dupuis) and Disco (JL Avila, C. Bonnet), and J.-P. Marie's team at St Antoine Hospital leukaemic tumour bank, where A. Ballesta, Cancéropole IdF-Inria postdoc has been detached (ending in March 2013) to identify parameters of a model of acute myeloblastic leukaemia (AML) in patient fresh cell cultures with and without anticancer drugs. This work has led to several presentations, and publications are in preparation.
3. *Hybrid models* Systems combining PDEs and discrete representations in hybrid models, with applications to cancer growth and therapy, in particular for AML, are the object of study of the ANR program *Bimod*, coordinated by V. Volpert (Lyon), associating CNRS (V. Volpert, Lyon), Bordeaux II University (P. Magal) and the Bang project-team.
4. *Molecular model of the activity of the p53 protein.* This work, the object of Luna Dimitrio's PhD thesis [1], co-supervised by J. Clairambault and R. Natalini (Rome), has led to her PhD defence in September 2012 at UPMC, and to a first publication [18], that should be followed by others. After L. Dimitrio's leave for the pharmaceutic industry, a new PhD student, Ján Eliš, has taken over this theme in September 2012 in a new PhD thesis at UPMC, under the supervision of J. Clairambault and B. Perthame

6.1.2. Physiological and pharmacological control of cell proliferation

Participants: Annabelle Ballesta, Frédérique Billy, Jean Clairambault, Sandrine Dulong [INSERM Villejuif (U 776)], Olivier Fercoq [MaxPlus project-team], Stéphane Gaubert [MaxPlus project-team], Thomas Lepoutre [Dracula project-team], Francis Lévi [INSERM Villejuif (U 776)].

1. *Periodic (circadian) control of cell proliferation in a theoretical model of the McKendrick type.* This theme (cf. supra "transition kernels...") has been continued [39], [11], [12], [10], [42], [43]. Whereas transition kernels between cell cycle phases without control have been experimentally identified in cell cultures by FUCCI imaging [12], their circadian control remains elusive and has been modelled on the basis of gating by plain cosines representing the influence exerted on these transition kernels by circadian clocks. To go further, it would be necessary to have access by cell imaging to the activity of the best physiological candidates to such gating, namely the cyclin-Cdk complexes, together with the activities of the clock-controlled proteins Wee1 and p21, which thus far have remained unavailable to us through biological experimentation with imaging.
2. *Intracellular pharmacokinetic-pharmacodynamic (PK-PD) models for anticancer drugs.* This theme has continued to be developed with new publications for the drugs irinotecan [40], [44], 5-fluorouracil and oxaliplatin [43].

6.1.3. Optimisation of cancer chemotherapy

Participants: Annabelle Ballesta, Frédérique Billy, Frédéric Bonnans [Commands project-team], Jean Clairambault, Sandrine Dulong [INSERM Villejuif (U 776)], Xavier Dupuis [Commands project-team], Olivier Fercoq [MaxPlus project-team], Stéphane Gaubert [MaxPlus project-team], Thomas Lepoutre [Dracula project-team], Alexander Lorz, Francis Lévi [INSERM U 776, Villejuif], Michael Hochberg [ISEM, CNRS, Montpellier], Benoît Perthame.

Optimising cancer chemotherapy, in particular chronotherapy, is the final aim of the activities mentioned above. This theoretical activity has been continued, using the McKendrick paradigm in works involving the *C5Sys* network [12], [42], [43], with numerical optimisation algorithms for the toxicity constraint, and also in more general settings taking into account another major issue of anticancer treatment, namely resistance to drugs in cancer cells. To this latter aim, we have developed another type of models based on integro-differential equations, which are inspired from those used in ecology for Darwinian evolution. These are aimed at studying another major issue in cancer therapy: appearance of resistances to treatment in tumour cell populations. Indeed, these cell populations, because of their heterogeneity and genomic instability, present an ability to adapt and evolve (in the Darwinian sense) that is much higher than in healthy cell populations [10], [27], [39]. The time scales under investigation, much shorter than in ecology, are still much longer than in microbiology, and are those of clinical treatments.

From a molecular point of view, studying drug resistance leads to the study of ABC transporters, which is one of the tracks followed by A. Ballesta, following her PhD thesis, in collaboration with F. Lévi's INSERM team in Villejuif [40], [44].

Underway is also the use of methods of optimal control developed by the Commands project-team (F. Bonnans, X. Dupuis) to optimise therapies in the treatment of Acute Myeloblastic Leukaemia (AML, cf. supra "Modelling haematopoiesis with applications to AML").

6.1.4. Protein polymerisation and application to amyloid diseases (ANR grant TOPPAZ)

Participants: Annabelle Ballesta, Vincent Calvez [ENS Lyon], Marie Doumic-Jauffret, Pierre Gabriel, Hadjer Wafaâ Haffaf, Benoît Perthame, Stéphanie Prigent [BPCP, INRA Jouy-en-Josas], Human Rezaei [BPCP, INRA Jouy-en-Josas], Léon Matar Tine [SIMPAF project-team, Inria Lille Nord-Europe].

Published in PLoS One in collaboration with the biologists' team of H. Rezaei [29], a new and very complete PDE model for protein polymerisation has been designed. Following F. Charles's work, A. Ballesta has applied this model to Huntington's disease (PolyQ expansion) and compared it with its ODE counterpart, leading to a better understanding of the leading mechanisms responsible for PolyQ fibrillisation. New applications of this framework model are in progress with H.W. Haffaf and S. Prigent.

The eigenvalue problem playing a major role in the representation of Prion proliferation dynamics and, in a more general way, of many fragmentation-coalescence phenomena, the article [15] published in J. de Math. Pur. Appl. investigated the dependency of the principal eigenvector and eigenvalue upon its parameters. We exhibited possible nonmonotonic dependency on the parameters, conversely to what would have been conjectured on the basis of some simple cases.

6.1.5. Inverse problem in growth-fragmentation equations

Participants: Marie Doumic-Jauffret, Marc Hoffmann [ENSAE], Nathalie Krell [Univ. Rennes I], Patricia Reynaud [CNRS, Nice Univ.], Lydia Robert [UPMC], Vincent Rivoirard [Paris IX Univ.], Léon Matar Tine [SIMPAF project-team, Inria Lille Nord-Europe].

In collaboration with statisticians (M. Hoffman, Professor at Université de Marne-la-Vallée, V. Rivoirard, MC at Université d'Orsay, and P. Reynaud, CR CNRS at Université de Nice), in the article [19] published in SIAM Num. Anal., we explored a statistical viewpoint on the cell division problem. In contrast to a deterministic inverse problem approach, we take the perspective of statistical inference. By estimating statistically each term of the eigenvalue problem and by suitably inverting a certain linear operator, we are able to construct an estimator of the division rate that achieves the same optimal error bound as in related deterministic inverse problems. Our procedure relies on kernel methods with automatic bandwidth selection. It is inspired by model selection and recent results of Goldenschluger and Lepski.

An extension of this work, which consists of the statistical estimation of a branching process modelling the same growth and fragmentation dynamics, has been submitted in [49], in collaboration with N. Krell, M. Hoffmann and L. Robert.

In [20], published in J. Math. Biol. with L. Matar Tine, we generalised the inverse techniques proposed previously in [53], [57], in order to adapt them to general fragmentation kernels and growth speeds. The potential applications of this problem are numerous, ranging from polymerisation processes to the cell division cycle. An extension of this work is in progress with M. Escobedo and T. Bourgeron.

6.2. Tissue growth, regeneration and cell movements

6.2.1. Chemotaxis, self-organisation of cell communities (KPP-Fisher and Keller-Segel)

Participants: Nikolaos Bournaveas [Univ. Edinburgh], Axel Buguin [UPMC, Institut Curie], Vincent Calvez [ENS Lyon], François James [univ. Orléans], Alexander Lorz, Grégoire Nadin [UPMC], Benoît Perthame, Jonathan Saragosti [Institut Curie], Pascal Silberzan [Institut Curie], Min Tang [Shanghai Jiaotong University], Nicolas Vauchelet.

Chemotaxis denotes the ability of some cells to undergo a directed movement in response to an extracellular chemical substance. A mathematical description of chemotaxis is a major issue in order to understand collective movements of bacterial colonies. Numerous mathematical models, at various scales, have been proposed, allowing for a good description of cell aggregation under chemotaxis at the macroscopic level, the first of all being that of Keller-Segel (1971), that is now at the centre of an abundant international scientific literature.

At the cell scale, one uses kinetic equations. Numerical simulations have been performed and blow-up is also observed, which differs highly from pointwise blow-up in parabolic models. Representing them leads to various theoretical questions and amounts to define measure solutions [25], [24] or to develop an existence theory.

6.2.2. Single-cell-based and continuum models of avascular tumours

Participants: Ibrahim Cheddadi, Dirk Drasdo, Benoît Perthame, Min Tang [Shanghai Jiaotong University], Nicolas Vauchelet, Irène Vignon-Clémentel [REO project-team].

The recent biomechanical theory of cancer growth considers solid tumours as liquid-like materials comprising elastic components. In this fluid mechanical view, the expansion ability of a solid tumour into a host tissue is mainly driven by either diffusion of cells (emerging on the mesoscopic scale by coarse graining from the cell micro-motility) or by cell division depending either on the local cell density (contact inhibition), on mechanical stress in the tumour, or both. For the two by two degenerate parabolic/elliptic reaction-diffusion system that results from this modelling, we prove there are always travelling waves above a minimal speed and we analyse their shapes. They appear to be complex with composite shapes and discontinuities. Several small parameters allow for analytical solutions; in particular the incompressible cells limit is very singular and related to the Hele-Shaw equation. These singular travelling waves are recovered numerically. See [32].

6.2.3. Single cell-based models of tumour growth, tissue regeneration

Participants: Gregory Batt [CONTRAINTES project-team], François Bertaux, Géraldine Cellière, Chadha Chettaoui, Ibrahim Cheddadi, Dirk Drasdo, Adrian Friebel, Rolf Gebhardt [Univ. of Leipzig, Germany], Adriano Henney [Director Virtual Liver Network and VLN consortium], Jan G. Hengstler [Leibniz Research Centre, Dortmund, Germany and CANCERSYS consortium], Stefan Höhme, Elmar Heinzle [University of Saarbrücken and NOTOX consortium], Isabelle Hue [INRA], Nick Jagiella, Ursula Klingmüller [German Cancer Centre, Heidelberg and LungSys Consortium], Axel Krinner, Johannes Neitsch, Benoît Perthame, Ignacio Ramis-Conde, Luc Soler [IRCAD, Coordinator EU-project PASSPORT and PASSPORT consortium], Jens Timmer [University of Leipzig, Germany], Irène Vignon-Clémentel [REO project-team], Juhui Wang [INRA], William Weens.

6.2.3.1. A Multi-scale model for clonal competition in growing tumours

In this work we set up a multi-scale model testing the impact of three experimentally found variants of a signal transduction pathway controlling cell-cell adhesion on multi-cellular growth as well as the possible consequences of inhomogeneous populations where each of the three phenotypes competed [30].

6.2.3.2. Growth of cell populations in embedding granular and cell-like matter

In this work simulations of growing 2D and 3D clones embedded in granular and cell-like matter were mimicked [21]. The influence of active directed cell motion vs. passive pushing triggered by cell proliferation, as well as of various parameters of the embedding matter, such as the friction of embedding objects with its environment, adhesion strength, size of objects, elastic modulus etc. on the growth kinetics and the spatial pattern has been studied. The emerging patterns are strongly reminiscent of a fingering instability (a type of a Saffman-Taylor instability) occuring if a viscous fluid is injected into a more viscous fluid constrained between two plates (Hele Shaw cell).

6.2.3.3. Quantitative modelling of multi-cellular spheroids

Nick Jagiella in his thesis has worked out how stepwise and iteratively mechanisms controlling the spatial-temporal growth dynamics can be inferred by combining information from bright field micrographs stained for proliferating, dying cells, cell nuclei and extra-cellular matrix with the macroscopic growth kinetics.
This thesis, pursued within the German network project LUNGSYS was defended in September 2012. The thesis work was mainly supervised by Dirk Drasdo, PI for this part within the LUNGSYS project. Main collaborators were Margareta Mueller (previously DKFZ, Heidelberg) and Ursula Klingmueller, (DKFZ Heidelberg).
Moreover, Géraldine Cellière has worked out a model to mimic the aggregation of cells in the hanging drop method, a standard method to generate 3D multi-cellular aggregates. The kinetics and final configuration give information on multicellular aggregates. This work is pursued within the EU NOTOX project. Main collaborators are Fozia Noor and Elmar Heinzle (Univ. of Saarbruecken).

6.2.3.4. Image reconstruction of 3D liver architecture at subcellular level

In order to permit simulation liver function we started to set up an image processing pipeline resolving liver at subcellular scale. This will enable us to mimic all flows in liver, which comprises of blood flow through the micro-vessels (sinusoids), of blood plasma through the space between micro-vessel wall and hepatocytes, the main type of liver cells (called space of Disse), and of the bile through a network of bile canaliculi. Besides image analysis, also setting up the models of the flows has been started.

This work is conducted by the PhD student Adrian Friebel (IZBI, University of Leipzig) co-supervised by Dirk Drasdo and Stefan Hoehme (IZBI, University of Leipzig) within the Germany funded grant project Virtual Liver Network (VLN; PI from IZBI, Leipzig: Dirk Drasdo). Main collaborator is Jan G. Hengstler from the IfADo (directeur at the Leibniz Institute in Dortmund, Germany).

6.2.3.5. Ammonia metabolism during liver regeneration

Based upon the paper on liver regeneration after drug-induced damage (Hoehme et. al. PNAS 2010 [55]) we in a next step investigated the change of ammonia metabolism during the regeneration process. Ammonia is toxic for the body. We linked our spatial-temporal liver lobule model with a compartment model for the ammonia, glutamine and urea metabolism. In the latter we consider a compartment (the peri-central compartment) in which glutamine synthetase, a strongly ammonia-detoxifying enzyme, is degraded efficiently and a (peri-portal) compartment, in which this is not the case. By testing different hypotheses on the chemical reactions taking place during the degradation process and quantitatively comparing to time-space data of the regeneration process including data on the activity of glutamine synthetase we were able to propose a potentially missing chemical reaction. Validation experiments have been started and suggest that the original reaction scheme was indeed incomplete.

This work is conducted by Dirk Drasdo and Stefan Hoehme (IZBI, University of Leipzig) partly within the Germany funded grant project Virtual Liver Network (VLN; PI from IZBI, Leipzig: Dirk Drasdo) and the EU project NOTOX. Main collaborators are Rolf Gebhardt (chair for Biochemistry, University of Leipzig), Jan G. Hengstler from the IfADo (Leibniz Institute in Dortmund, Germany) and BioControl Jena GmbH, a company in Jena, Germany.

6.2.3.6. Multi-scale simulation of cell cycle progression during liver regeneration

In previous work on liver regeneration after drug induced damage (Hoehme et. al. PNAS 2010 [55]) the experimentally observed spatial-temporal proliferation pattern has been used as an input parameter. We have now started to study the molecular control of cell cycle progression by hepatocyte growth factor (HGF). Based on model predictions with a hypothesized model linking the downstream activation of the HGF-pathway with cell cycle progression, experiments were performed which now led to a validated intracellular model of cell cycle progression by HGF. Moreover, based on model simulations predicting that two sources of HGF are necessary to explain the experimentally observed proliferation pattern, experiments detecting the potential sources of HGF have been initiated. The models are multi-scale in that the precise spatial architecture of a piece of liver tissue is modelled representing each individual hepatocyte as well as the blood micro-vessels. A system of ODE's mimicking the HGF signalling and its impact on cell cycle progression is solved inside each individual cell. The project works out a systematic strategy to stepwise identify multi-scale multi-level processes in tissue organisation extending the lines pursued in Hoehme et. al. [55] and Holzhuetter et. al. [23].
This work is conducted by Dirk Drasdo and Stefan Hoehme (IZBI, University of Leipzig) within the Germany funded grant project Virtual Liver Network (VLN; PI from IZBI, Leipzig: Dirk Drasdo). Main collaborators are Ursula Klingmueller and Lorenza D'Alessandro (UK is Professor at Heidelberg University and department head at German Cancer Research Centre (DKFZ), Heidelberg, Germany) as well as Jens Timmer and Andreas Raue (JT is Professor University of Freiburg, Germany).

6.2.3.7. Phenotypes in early liver cancer

The model of a liver lobule, the smallest functional unit of liver (Hoehme et. al., PNAS 2010 [55]) has been used as a starting point to explain the experimentally observed early tumour phenotypes. We made a sensitivity analysis to identify the parameters that influence the tumour phenotype. Each simulation mimicked a monoclonal tumour. We could show that the observed early phenotypes could be explained by only a few sensitive parameters which are the direction of cell division, cell-micro-vessel adhesion, and destruction of micro-vessels by the tumour cells.
This work has been taken over from the previous PhD student William Weens by the PhD student François Bertaux who is co-supervised by Dirk Drasdo and Gregory Batt. Main collaborator is Jan G. Hengstler from the IfADo (directeur at the Leibniz Institute in Dortmund, Germany).

6.2.3.8. Regeneration of liver after partial hepatectomy

We continued this earlier activity by initiating experiments on pigs to test the model prediction that the 2nd wave of proliferation during regeneration after partial hepatectomy in pig should occur only close to the Glisson capsule, that encloses the liver, while in mouse proliferation occurs homogeneously and isotropically distributed over the whole liver lobe.
This work is conducted by Dirk Drasdo and Stefan Hoehme (IZBI, University of Leipzig) within the Germany funded grant project Virtual Liver Network. Main collaborators are Jan G. Hengstler from the IfADo (Leibniz Institute in Dortmund, Germany) and Eric Vilbert, Centre Hépato-Biliaire (CHB)- INSERM U785, Hospital Paul Brousse, Villejuif.

6.2.3.9. High resolution model for eukaryotic cells

In order to permit simulations directly out of 3D reconstructions of confocal laser scanning micrographs at subcellular resolution we developed a model that is capable to resolve complex cell shapes. The model parameters were calibrated by comparison with experiments probing the material properties of cells. Moreover, the cell division was implemented. The model was integrated into the CellSys software.
This work is conducted by the PhD student Johannes Neitsch (IZBI, University of Leipzig) co-supervised by Dirk Drasdo and Stefan Hoehme (IZBI, University of Leipzig) within the Germany funded grant project Virtual Liver Network (VLN). Main collaborators are Jan G. Hengstler from the IfADo (directeur at the Leibniz Institute in Dortmund, Germany) and Josef Kaes (Prof. for Experimental Physics, Univ. Leipzig).

6.2.3.10. Yeast cells playing the Game of Life

Within a collaboration with a synthetic biology lab at MIT, we work on the multicellular modelling of engineered yeast cell populations. Those cells secrete a messenger molecule (IP) which diffuse in the medium, bind to other cells, and trigger a signalling cascade which finally induce expression of lethal genes. A model has been established based on our single-cell-based model framework associated with PDE's simulations, and it is currently used to explain and guide experiments obtained at MIT.
This work is conducted within the project Sine2Arti by François Bertaux co-supervised by Gregory Batt and Dirk Drasdo, and by Szymon Stoma. Main collaborator is Ron Weiss, MIT, Boston, USA.

6.2.3.11. Stochastic modelling of extrinsic apoptosis

Here we extended a well-established ODE model of TRAIL-induced apoptosis developed by Sorger's group in Harvard by the possible effect of cell-to-cell variability due to stochasticity of rare events in the cascade.
This work is conducted within the project Sine2Arti by François Bertaux co-supervised by Gregory Batt and Dirk Drasdo, and by Szymon Stoma as well as Xavier Duportet for the experimental part.

6.2.3.12. Artificial Homeostasis in HeLa cells

The aim is to genetically engineer human cancer cells (HeLa cell line) such that they perform population control in a petri dish. To do so, it is made use of extrinsic apoptosis by forcing cells to produce a messenger molecule able to trigger apoptosis above a certain threshold concentrations in the medium. We developed a mathematical model which integrates both PDEs and intracellular components into a single-cell-based model framework. Such model allows to help designing the genetic system that should be integrated into cells as well as guiding experiments.
This work is conducted within the project Sine2Arti by François Bertaux who is co-supervised by Gregory Batt and Dirk Drasdo. Moreover Szymon Stoma for the modelling part, as well as Xavier Duportet for the experimental part from the CONTRAINTES team are included.

6.2.4. Modelling flow in tissues

Participants: Lutz Brusch [TU Dresden], Dirk Drasdo, Adrian Friebel [IZBI, University of Leipzig], Stefan Hoehme [IZBI, University of Leipzig], Nick Jagiella [Inria and IZBI, University of Leipzig], Hans-Ulrich Kauczor [University of Heidelberg, Germany], Fabian Kiessling [University Clinics, Technical University of Aachen, Germany], Ursula Klingmueller [German Cancer Research Centre (DKFZ), Heidelberg, Germany], Hendrik Laue [Fraunhofer Mevis, Bremen, Germany], Ivo Sbazarini [MPI for Molecular Cell Biology and Genetics, Dresden, Germany], Irène Vignon-Clémentel [REO project-team], Marino Zerial [MPI for Molecular Cell Biology and Genetics, Dresden, Germany].

6.2.4.1. Flow and perfusion scenarios in cancer

In this subject we simulated typical flow and perfusion scenarios in tumour and tissue including, how the spatial-temporal pattern look like on the scale of non-invasive medical image modalities currently applied, to infer parameters that are used to or may permit to evaluate the perfusion of tumors in patients. The simulations use Poiseuille flow and Kirchhoff rule in 3D blood network representing typical architectures.
The work was part of the PhD thesis of Nick Jagiella, defended in September 2012 co-supervised by Dirk Drasdo and Irene Vignon-Clementel, and conducted within the grant funded network projects LUNGSYS and LUNGSYS II. Main collaborators were Oliver Sedlaczek, DKFZ Heidelberg and University of Heidelberg, Fabian Kissling, Technical University of Aachen and Hendrik Laue, Fraunhofer Mevis, Bremen (all in Germany).

6.2.4.2. Flow in liver lobules

The aim of this project is to simulate realistically the flow of matter within liver lobules from images generated with different image modalities at histological scales. So far we have established a model of blood flow and perfusion in liver lobules based upon 3D reconstruction of confocal micrographs.
This work is conducted by collaboration of different groups within the Germany funded grant project Virtual Liver Network. From our group Nick Jagiella, Adrian Friebel, and Stefan Hoehme, Dirk Drasdo are involved, main collaborators are Irene Vignon-Clementel (REO project team Inria), Marino Zerial and Ivo Sbazarini (Max-Planck Institute for Molecular Cell Biology and Genetics, Dresden, Germany), Lutz Brusch (Technical University of Dresden) and Jan G. Hengstler from the IfADo (Leibniz Institute in Dortmund, Germany).

6.2.5. Contraction of acto-myosin structures in morphogenesis and tissue repair

Participants: Luís Almeida, P. Bagnerini [Univ. Genova], A. Habbal [Univ. Nice], A. Jacinto [CEDOC, Lisbon], M. Novaga [Univ. Padova], A. Chambolle [École Polytechnique], J. Demongeot [Univ. Grenoble].

Contraction of actin structures (in one, two or three dimensions) plays an important role in many cellular and tissue movements, both at a multicellular tissue level and at a cellular (and even intracellular) one: from muscle contraction to neural tube closure, epiboly in zebrafish embryo, the contractile ring in cytokinesis, cell crawling,... examples are everywhere in the living world. These structures consist of meshworks of actin filaments (which are like fibers) that are cross-linked by molecular motors (Myosin II) which can make the actin filaments slide relative to each other, thus generating deformation movements.

In [4] we are particularly interested in modelling the contraction of acto-myosin cables in morphogenesis and tissue repair. The experiments done in collaboration with A. Jacinto's lab show that the local curvature (and in particular its sign) plays an important role in the contractile behaviour of the acto-myosin cables. These experimental results led us to develop some of these ideas in [6] and to do a more abstract study of flows by the positive part of the curvature in [5].

6.3. Neurosciences

Participants: M. Galtier, G. Hermann, M. Magnasco, T. Taillefumier, Jonathan Touboul.

We pursued the analysis of the dynamics of networks of neurons in the presence of noise. Limit theorems in simple cases were treated in [9], and more refined models including space, delays and heterogeneities were analysed in [34], [35], touboulde lays:12,touboulNeuralFieldsDynamics:12. In all these contributions we analysed the eminently important role of noise and heterogeneity on the qualitative dynamics of networks. Mathematical results were obtained for representation of the solutions to linear functional differential equations [22] that were motivated by plasticity phenomena in the cortex.

6.4. Free surface geophysical flows

Participants: Emmanuel Audusse [LAGA - Université Paris 13, Institut Galilée], Anne-Céline Boulanger, Marie-Odile Bristeau, Benoît Perthame, Jacques Sainte-Marie, Nicolas Seguin, Edwige Godlewski, Anne Mangeney, Yohan Penel, Raouf Hamouda, Philippe Ung.

The ANGE team has been created in november 2012. This new team (led by J. Sainte-Marie) resumes the activities of the BANG team concerning geophysical flows.

We are involved in research concerning the numerical simulation of free surface geophysical flows such as rivers, lakes, coastal areas and also overland flows. Many applications related to environmental problems are concerned : floodings, dam breaks, swell, transport and diffusion of pollutants, water quality, upwellings, sustainability of aquatic ecosystems, ...

The basic model for these problems is the 3D free surface Stokes system leading to a 3D solver [52] with a moving mesh. However for efficiency reasons, vertically averaged models such as the Saint-Venant system [54] are often used.

The Saint-Venant equations are deduced of the Navier-Stokes system with two main assumptions:

- the pressure is hydrostatic,
- the horizontal velocity is represented by its average.

We have developed extensions of the Saint-Venant system where the basic Saint-Venant solver [51] is still used and, in that way, the robustness, the efficiency and the easiness to treat the free surface are preserved while the domain of validity is larger.

In these extensions, we relax the two above assumptions. Actually, we have derived a non-hydrostatic shallow water model and a multilayer Saint-Venant system.

We have coupled the hydrodynamics of free surface flows with other phenomena such as biology (phytoplankton culture) or erosion.

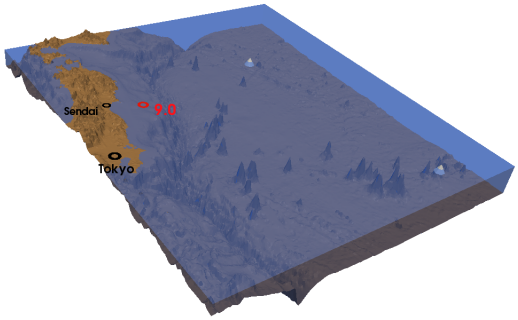

Figure 1. Map of Japan with the seism epicentre and the DART buoys 21418 and 21413.

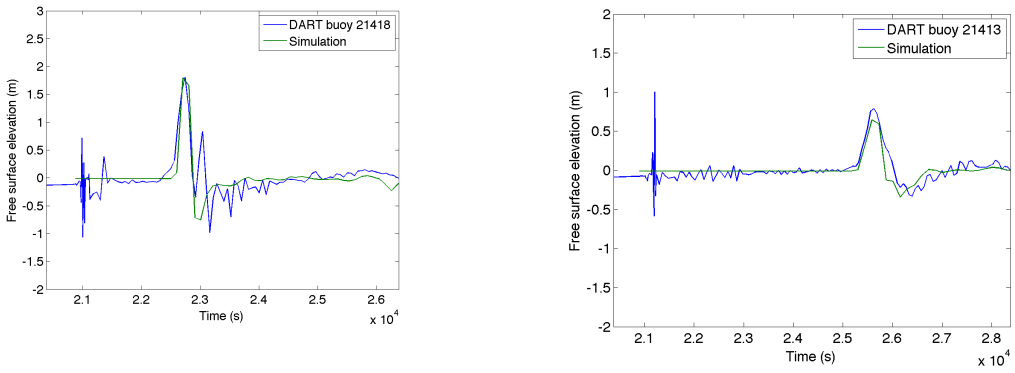

Figure 2. Free surface elevation of the sea, comparison between the recorded data by the buoys 21418 and 21413 and the simulation obtained with our 3d Navier-Stokes code.

CLIME Project-Team

6. New Results

6.1. New methods for data assimilation

Since the beginning, Clime is focused on developing new techniques for data assimilation in geophysical sciences. Clime is active on several of the most challenging theoretical aspects of data assimilation: data assimilation methods based on non-Gaussian assumptions, methods for estimating errors, ensemble filtering techniques, 4D variational assimilation approaches, ensemble-variational methods, etc. This year, we revisited several of these topics. A dual algorithm has been developed for the finite-size ensemble Kalman filter, that shows how to estimate optimal inflation that counteracts sampling errors. A variational method coupled to a subgrid scale statistical model has been introduced and validated to quantify the representativeness errors. We also started to work on ensemble variational methods that are brand new techniques emerging in the meteorological data assimilation field.

6.1.1. Combining inflation-free and iterative ensemble Kalman filters for strongly nonlinear systems

Participants: Marc Bocquet, Pavel Sakov [NERSC, Norway].

The finite-size ensemble Kalman filter (EnKF-N) is an ensemble Kalman filter (EnKF) which, in perfect model condition, does not require inflation because it partially accounts for the ensemble sampling errors. For the Lorenz '63 and '95 toy-models, it was so far shown to perform as well or better than the EnKF with an optimally tuned inflation. The iterative ensemble Kalman filter (IEnKF) is an EnKF which was shown to perform much better than the EnKF in strongly nonlinear conditions, such as with the Lorenz '63 and '95 models, at the cost of iteratively updating the trajectories of the ensemble members. This study aims at further exploring the two filters, and at combining both into an EnKF that does not require inflation in perfect model condition and which is as efficient as the IEnKF in very nonlinear conditions.

In this study EnKF-N is first introduced and a new implementation is developed. It decomposes EnKF-N into a cheap two-step algorithm that amounts to computing an optimal inflation factor. This offers a justification of the use of the inflation technique in the traditional EnKF and why it can often be efficient. Secondly, the IEnKF is introduced following a new implementation based on the Levenberg-Marquardt optimization algorithm. Then, the two approaches are combined to obtain the finite-size iterative ensemble Kalman filter (IEnKF-N). Several numerical experiments are performed on IEnKF-N with the Lorenz '95 model. These experiments demonstrate its numerical efficiency as well as its performance that offer, at least, the best of both filters.

6.1.2. Accounting for representativeness errors in the inversion of atmospheric constituent emissions: Application to the retrieval of regional carbon monoxide fluxes

Participants: Mohammad Reza Koohkan, Marc Bocquet.

A four-dimensional variational data assimilation system (4D-Var) is developed to retrieve carbon monoxide (CO) fluxes at regional scale, using an air quality network. The air quality stations that monitor CO are proximity stations located close to industrial, urban or traffic sources. The mismatch between the coarsely discretized Eulerian transport model and the observations, inferred to be mainly due to representativeness errors in this context, leads to a bias (averaged simulated concentrations minus observed concentrations) of the same order of magnitude as the concentrations. 4D-Var leads to a mild improvement in the bias because it does not adequately handle the representativeness issue. For this reason, a simple statistical subgrid model is introduced and is coupled to 4D-Var. In addition to CO fluxes, the optimization seeks to jointly retrieve *influence coefficients*, which quantify each station's representativeness. The method leads to a much better representation of the CO concentration variability, with a significant improvement of statistical indicators.

The resulting increase in the total inventory estimate is close to the one obtained from remote sensing data assimilation. This methodology and experiments suggest that information useful at coarse scales can be better extracted from atmospheric constituent observations strongly impacted by representativeness errors.

6.1.3. Real-time data assimilation

Participants: Vivien Mallet, Anne Tilloy, Fabien Brocheton [Numtech], David Poulet [Numtech], Cécile Honoré [Airparif], Édouard Debry [INERIS].

Based on Verdandi, Polyphemus and the "Urban Air Quality Analysis" software, real-time data assimilation was carried out at urban scale. The Best Linear Unbiased Estimator (BLUE) was computed for every hourly concentration map that the ADMS model computed. A posteriori tests were conducted over Clermont-Ferrand and Paris. We addressed the key issue of the covariance of the state error. The form of the error covariance between two points was determined based on the road network, considering the distance between points along the road and the distance of each point to the road. A few parameters (primarily two decorrelation lengths) were determined thanks to cross validation with several months of simulations and observations. The results showed strong improvements even at locations where no data was assimilated.

At larger scale, the data assimilation library Verdandi was used to apply data assimilation (optimal interpolation) with the air quality model Chimere. This preliminary work will help INERIS to apply optimal interpolation for ozone and particulate matter in the operational platform Prev'air.

6.2. Inverse modeling

Many of this year's studies have focused on inverse modeling, including the reconstruction of the Fukushima radionuclide atmospheric and marine source terms. All were targeted to a particular application. However most of them include new methodological developments, in particular non-Gaussian data assimilation schemes.

6.2.1. Estimation of errors in the inverse modeling of accidental release of atmospheric pollutant: Application to the reconstruction of the Fukushima Daiichi source term

Participants: Victor Winiarek, Marc Bocquet, Olivier Saunier [IRSN], Anne Mathieu [IRSN].

The aim of this research activity is the implementation of data assimilation methods, particularly inverse modeling methods, in the context of an accidental radiological release from a nuclear power plant and their application in the specific case of the Fukushima Daiichi accident. The particular methodological focus is the a posteriori estimation of the prior errors statistics. In the case of the Fukushima Daiichi accident, the number of available observations is small compared to the number of source parameters to retrieve and the reconstructed source is highly sensitive to the prior errors. That is the why they need to be well established and justified. In this aim, three methods have been proposed: one method relies on a L-curve estimation technique, another one on the Desroziers' iterative scheme and the last method, assumed to be the most robust, relies on the maximum likelihood principle, generalised to a non-Gaussian context. These three methods have been applied to the reconstruction of cesium-137 and iodine-131 source terms from the Fukushima Daiichi accident. Because of the poor observability of the Fukushima Daiichi emissions, these methods provide lower-bounds for cesium-137 and iodine-131 reconstructed activities. Nevertheless, with the new method based on semi-Gaussian statistics for the background errors, the lower-bound estimates for cesium-137, $1.2 - 4.0\,10^{16}$ Bq with an estimated standard deviation range of $15 - 20\%$, and for iodine-131, $1.9 - 3.8\,10^{17}$ Bq with an estimated standard deviation range of $5 - 10\%$, are of the same order of magnitude as those provided by the Japanese Nuclear and Industrial Safety Agency, and about 5 to 10 times less than the Chernobyl atmospheric releases.

6.2.2. Assessment of the amount of Cesium-137 released into the Pacific Ocean after the Fukushima accident and analysis of its dispersion in Japanese coastal waters

Participants: Claude Estournel [LA], Emmanuel Bosc [IAEA], Marc Bocquet, Caroline Ulses [LA], Patrick Marsailex [LA], Victor Winiarek, Iolanda Osvath [IAEA], Cyril Nguyen [LA,LEGOS], Thomas Duhaut [LA], Florent Lyard [LEGOS], Héloïse Michaud [LA], Francis Auclair [LA].

Numerical modeling was used to provide a new estimate of the amount of cesium-137 released directly into the ocean from the Fukushima Daiichi nuclear power plant (NPP) after the accident in March 2011 and to gain insights into the physical processes that led to its dispersion in the marine environment during the months following the accident. An inverse method was used to determine the time-dependent cesium-137 input responsible for the observed concentrations. The method was then validated through comparisons of the simulated concentrations with concentrations measured in seawater at different points in the neighborhood of the nuclear power plant. An underestimation was noticed for stations located 30 km offshore. The resulting bias in the release inventory was estimated. Finally, the maximum cesium-137 activity released directly to the ocean was estimated to lie between 5.1 and 5.5 PBq (Peta Becquerel = 1015 Bq) but uncertainties remain on the amount of radionuclides released during the first few days after the accident. This estimate was compared to previous ones and differences were further analysed. The temporal and spatial variations of the cesium-137 concentration present in the coastal waters were shown to be strongly related to the wind intensity and direction. During the first month after the accident, winds blowing toward the south confined the radionuclides directly released into the ocean to a narrow coastal band. Afterwards, frequent northward wind events increased the dispersion over the whole continental shelf, leading to strongly reduced concentrations.

6.2.3. What eddy-covariance measurements tell us about prior land flux errors in CO2-flux inversion schemes?

Participants: Frédéric Chevallier [LSCE], Tao Wang [LSCE], Philippe Ciais [LSCE], Marc Bocquet, Altaf Arain [McMaster University, Canada], Alessandro Cescatti [Joint Research Centre, Italy], Jiquan Chen [University of Toledo, USA], Johannes Dolman [Vrije Universiteit, the Netherlands], Beverly Law [Oregon State University, USA], Hank Margolis [Université Laval, Canada], Leonardo Montagnani [University of Bolzano, Italy].

To guide the future development of CO2-atmospheric inversion modeling systems, we analysed the errors arising from prior information about terrestrial ecosystem fluxes. We compared the surface fluxes calculated by a process-based terrestrial ecosystem model with daily averages of CO2 flux measurements at 156 sites across the world in the FLUXNET network. At the daily scale, the standard deviation of the model-data fit was 2.5 gC·m-2·d-1; temporal autocorrelations were significant at the weekly scale (> 0.3 for lags less than four weeks), while spatial correlations were confined to within the first few hundred kilometers (< 0.2 after 200 km). Separating out the plant functional types did not increase the spatial correlations, except for the deciduous broad-leaved forests. Using the statistics of the flux measurements as a proxy for the statistics of the prior flux errors was shown not to be a viable approach. A statistical model allowed us to upscale the site-level flux error statistics to the coarser spatial and temporal resolutions used in regional or global models. This approach allowed us to quantify how aggregation reduces error variances, while increasing correlations. As an example, for a typical inversion of grid point (300 km $\times$ 300 km) monthly fluxes, we found that the prior flux error follows an approximate e-folding correlation length of 500 km only, with correlations from one month to the next as large as 0.6.

6.3. Monitoring network design

In this section, we report studies that are related to the evaluation of monitoring networks and to new monitoring strategies. This year, network designs techniques have been applied to the inverse modeling of CO2 fluxes.

6.3.1. Network design for mesoscale inversions of CO2 sources and sinks

Participants: Thomas Lauvaux [Pennsylvania State University, USA], Andy Schuh [Colorado State University, USA], Marc Bocquet, Lin Wu, Scott Richardson [Pennsylvania State University, USA], Natasha Miles [Pennsylvania State University, USA], Ken Davis [Pennsylvania State University, USA].

Recent instrumental deployments of regional observation networks of atmospheric CO2 mixing ratios have been used to constrain carbon sources and sinks using inversion methodologies. In this study, we performed sensitivity experiments using observation sites from the Mid Continent Intensive experiment to evaluate the required spatial density and locations of CO2 concentration towers based on flux corrections and error reduction analysis. In addition, we investigated the impact of prior flux error structures with different correlation lengths and biome information. We show that, while the regional carbon balance converged to similar annual estimates using only two concentration towers over the region, additional sites were necessary to retrieve the spatial flux distribution of our reference case (using the entire network of eight towers). Local flux corrections required the presence of observation sites in their vicinity, suggesting that each tower was only able to retrieve major corrections within a hundred of kilometers around, despite the introduction of spatial correlation lengths (100 to 300 km) in the prior flux errors. We then quantified and evaluated the impact of the spatial correlations in the prior flux errors by estimating the improvement in the CO2 model-data mismatch of the towers not included in the inversion. The overall gain across the domain increased with the correlation length, up to 300 km, including both biome-related and non-biome-related structures. However, the spatial variability at smaller scales was not improved. We conclude that the placement of observation towers around major sources and sinks is critical for regional-scale inversions in order to obtain reliable flux distributions in space. Sparser networks seem sufficient to assess the overall regional carbon budget with the support of flux error correlations, indicating that regional signals can be recovered using hourly mixing ratios. However, the smaller spatial structures in the posterior fluxes are highly constrained by assumed prior flux error correlation lengths, with no significant improvement at only a few hundreds of kilometers away from the observation sites.

6.3.2. Potential of the International Monitoring System radionuclide network for inverse modeling

Participants: Mohammad Reza Koohkan, Marc Bocquet, Lin Wu, Monika Krysta [The Preparatory Commission for the Comprehensive Nuclear Test-Ban Treaty Organization, UNO].

The International Monitoring System (IMS) radionuclide network enforces the Comprehensive Nuclear-Test-Ban Treaty, which bans nuclear explosions. We have evaluated the potential of the IMS radionuclide network for inverse modeling of the source, whereas it is usually assessed by its detection capability. To do so, we have chosen the *degrees of freedom for the signal* (DFS), a well established criterion in remote sensing, in order to assess the performance of an inverse modeling system. Using a multiscale data assimilation technique, we have computed optimal adaptive grids of the source parameter space by maximizing the DFS. This optimization takes into account the monitoring network, the meteorology over one year (2009) and the relationships between the source parameters and the observations derived from the FLEXPART Lagrangian transport model. Areas of the domain, where the grid-cells of the optimal adaptive grid are large, emphasize zones where the retrieval is more uncertain, whereas areas, where the grid-cells are smaller and denser, stress regions where more source variables can be resolved. The observability of the globe through inverse modeling is studied in strong, realistic and small model error cases. The strong error and realistic error cases yield heterogeneous adaptive grids, indicating that information does not propagate far from the monitoring stations, whereas in the small error case, the grid is much more homogeneous.

In all cases, several specific continental regions remain poorly observed such as Africa as well as the tropics, because of the trade winds.

The northern hemisphere is better observed through inverse modeling (more than 60% of the total DFS), mostly because it contains more IMS stations. This unbalance leads to a better performance of inverse modeling in the northern hemisphere winter. The methodology is also applied to the subnetwork composed of the stations of the IMS network that measure noble gases.

6.4. Reduction and emulation

The use of environmental models raise a number of problems due to:

- the dimension of their inputs, which can easily be $10^5 - -10^8$ at every time step;

- the dimension of their state vector, which is usually $10^5 - -10^7$;
- their high computational cost.

In particular, the application of data assimilation methods and uncertainty quantification techniques may require dimension reduction and cost reduction. The dimension reduction consists in projecting the inputs and the state vector to low-dimensional subspaces. The cost reduction can be carried out by emulation, i.e., the replacement of costly components with fast surrogates.

6.4.1. Reduction and emulation of a chemistry-transport model

Participants: Vivien Mallet, Serge Guillas [University College London].

Both reduction and emulation were applied to the dynamic air quality model Polair3D from Polyphemus. The reduction relied on proper orthogonal decomposition on the input data and on the state vector. The dimension of the reduced subspace for the input data is about 80, while the dimension of the reduced state vector is less than 10. The projection of the state vector on its reduced subspace can be carried out before every integration time step, so that one can reproduce a full state trajectory (in time) using the reduced model.

Significant advances were made to emulate the reduced model, which requires about 90 inputs (reduced input data and reduced state vector) and computes about 10 outputs (reduced state vector). 90 inputs is however a large number to build an emulator using a classical approach like krigging. Promising results were however obtained with an interpolation method based on inverse distance weighting.

6.4.2. Reduction and emulation of a static air quality model

Participants: Vivien Mallet, Anne Tilloy, Fabien Brocheton [Numtech], David Poulet [Numtech].

The dimension reduction was applied to the outputs of the urban air quality model ADMS Urban, which is a static model with low-dimensional inputs and high-dimensional outputs. A proper orthogonal decomposition on the outputs allowed us to drastically reduce their dimension, from 10^4 to just a few scalars. First attempts of emulation of the reduced model rely on Gaussian process emulation.

6.4.3. Motion estimation from images with a wavelets reduced model

Participants: Giuseppe Papari, Isabelle Herlin, Etienne Huot, Karim Drifi.

The dimension reduction was applied to an image model, composed of Lagrangian constancy of velocity and transport of image brightness. Wavelets basis have been computed on the image domain for subspaces of images, motion fields and divergence-free motion fields. Image assimilation with this reduced model allows to estimate smooth velocity fields with properties defined by user. This also solves the issue of complex geographical domains and the difficulty of applying boundary conditions on these domains. First results are obtained with a reduced dimension of motion to a few scalars, to be compared with the original problem that has the size of image domain.

6.5. Ensemble forecasting with sequential aggregation

The aggregation of an ensemble of forecasts is an approach where the members of an ensemble are given a weight before every forecast time, and where the corresponding weighted linear combination of the forecasts provides an improved forecast. A robust aggregation can be carried out so as to guarantee that the aggregated forecast performs better, in the long run, than any linear combination of the ensemble members with time-independent weights. The approaches are then based on machine learning. The aggregation algorithms can be applied to forecast analyses (generated from a data assimilation system), so that the aggregated forecasts are naturally multivariate fields.

6.5.1. Application of sequential aggregation to meteorology and air quality

Participants: Anne Tilloy, Vivien Mallet, Fabien Brocheton [Numtech], David Poulet [Numtech].

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. Each time new observations are available, new weights for the linear combination are computed and applied for the next forecast. We applied the discounted ridge regression algorithm, which we previously introduced for sequential aggregation of air quality forecasts, to forecast wind and temperature at given observation stations. The ensemble was generated with forecasts at different range from two models. The aggregation proved to be efficient for one-day forecasts at least.

The discounted ridge regression was also applied to the simulations of the Air Quality Modeling Evaluation International Initiative (AQMEII) over Europe and North America, for different pollutants (gases and particulate matter).

6.5.2. Sequential aggregation with uncertainty estimation

Participants: Vivien Mallet, Sergiy Zhuk [IBM research], Paul Baudin, Gilles Stoltz [CLASSIC], Karine Sartelet [CEREA].

A new issue is the estimation of the uncertainties associated with the aggregated forecasts. One investigated direction relies on the framework of machine learning, with the aggregation of an ensemble of probability density functions instead of the point forecasts of the ensemble.

Another direction, which led to finalized results in 2012, is to reformulate the aggregation problem in a filtering problem for the weights. The weights are supposed to satisfy some dynamics with unknown model error, which defines the state equation of a filter. An observation equation compares the aggregated forecast with the observations (or analyses) with known observational error variance. The filter finally computes estimates for the weights and quantifies their uncertainties. We applied a Kalman filter and a minimax filter for air quality forecasting.

6.6. 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, these 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 problem is to control the location of structures using the observations,
- from the image's viewpoint, a model of the dynamics and structures has to be built from the observations.

6.6.1. Divergence-free motion estimation

Participants: Dominique Béréziat [UPMC], Isabelle Herlin, Sergiy Zhuk [IBM Research, Ireland].

This research addresses the issue of divergence-free motion estimation on an image sequence, acquired over a given temporal window. Unlike most state-of-the-art technics, which constrain the divergence to be small thanks to Tikhonov regularization terms, a method that imposes a null value of divergence of the estimated motion is defined.

Motion is characterized by its vorticity value and assumed to satisfy the Lagragian constancy hypothesis. An image model is then defined: the state vector includes the vorticity, whose evolution equation is derived from that of motion, and a pseudo-image that is transported by motion. An image assimilation method, based on the 4D-Var technics, is defined and developed that estimates motion as a compromise between the evolution equations of vorticity and pseudo-image and the observed sequence of images.

The method is applied on Sea Surface Temperature (SST) images acquired over Black Sea by NOAA-AVHRR sensors. The divergence-free assumption is roughly valid on these acquisitions, due to the small values of vertical velocity at the surface. Fig. 2 displays data and results. As no ground truth of motion is available, the method is quantified by the value of correlation between the pseudo-images and real acquisitions [28].

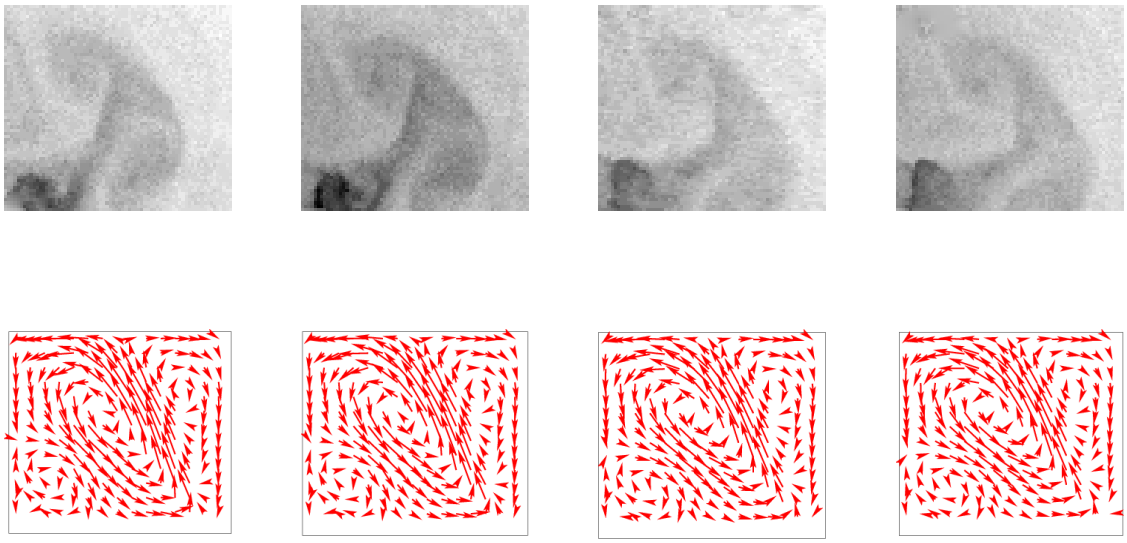

Figure 2. SST image observations and motion results.

6.6.2. Improvement of motion estimation by assessing errors on the dynamics

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, we consider adding a model error term in the evolution equation of motion and design a weak formulation of 4D-Var data assimilation. The cost function to be minimized simultaneously depends on the initial motion field, at the begining 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 [27].

This error assessment method is evaluated and quantified on twin experiments, as no ground truth would be available for real image data. Fig. 3 shows four frames of a series of observations obtained by integrating the evolution model from an initial condition on image and velocity field (the ground truth $\mathbf{w}_{\text{ref}}(0)$ displayed on the left of Fig. 4). An error value is added at each time step on the motion value, when integrating the simulation model. This error is a constant bias.

We performed two data assimilation experiments. The first one considers the evolution model as perfect, with no error in the evolution equation. It is denoted PM (for Perfect Model). The second one, denoted IM (for Imperfect Model) involves an error in the motion evolution equation. In Fig.4 are displayed the motion fields retrieved by PM and IM at the beginning of the temporal window.

As it can be seen, IM computes a correct velocity field while PM completely fails.

6.6.3. Nonlinear Observation Equation For Motion Estimation

Participants: Dominique Béréziat [UPMC], Isabelle Herlin.

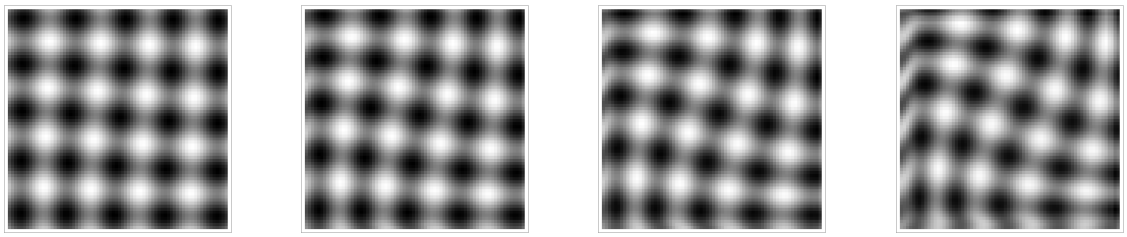

Figure 3. Observations Images.

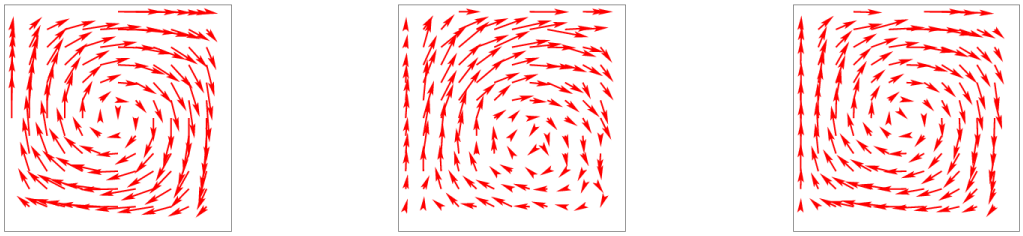

Figure 4. Left: ground-truth, middle: PM, right: IM.

In the image processing literature, the optical flow equation is usually chosen to assess motion from an image sequence. However, it corresponds to an approximation that is no more valid in case of large displacements. We evaluated improvements obtained when using the non linear transport equation of the image brightness by the velocity field [25]. A 4D-Var data assimilation method is designed that simultaneously solves the evolution equation and the observation equation, in its non linear and linearized form. The comparison of results obtained with both observation equations is quantified on synthetic data and discussed on oceanographic Sea Surface Temperature (SST) images. We show that the non linear model outperforms the linear one, which underestimates the motion norm. Fig.5 illustrates this on SST images (motion vectors are displayed by arrows).

The aim of this research is to achieve a correct estimation of motion when the object displacement is greater than its size. Howerver, in this case, coarse-to-fine incremental methods as well as the non linear data assimilation method fail to retrieve a correct value. The perspective is then to include, in the state vector, a variable describing the trajectory of pixels. The observation operator will then measure the effective displacement of pixels, according to their trajectories, and allow a better estimation of motion value.

6.6.4. Sliding windows method for motion estimation on long temporal image sequences

Participants: Karim Drifi, Isabelle Herlin.

This study concerns the estimation of motion fields from satellite images on long temporal sequences. The huge computational cost and memory required by data assimilation methods on the pixel grid makes impossible to use these techniques on long temporal intervals. For a given dynamic model (named full model), on the pixel grid, the Galerkin projection on subspaces provides a reduced model, that allows image assimilation at low cost. The definition of this reduced model however requires defining an optimal subspace of motion. A **sliding windows** method is thus designed:

- The long image sequence is split into small temporal windows that half overlap in time.

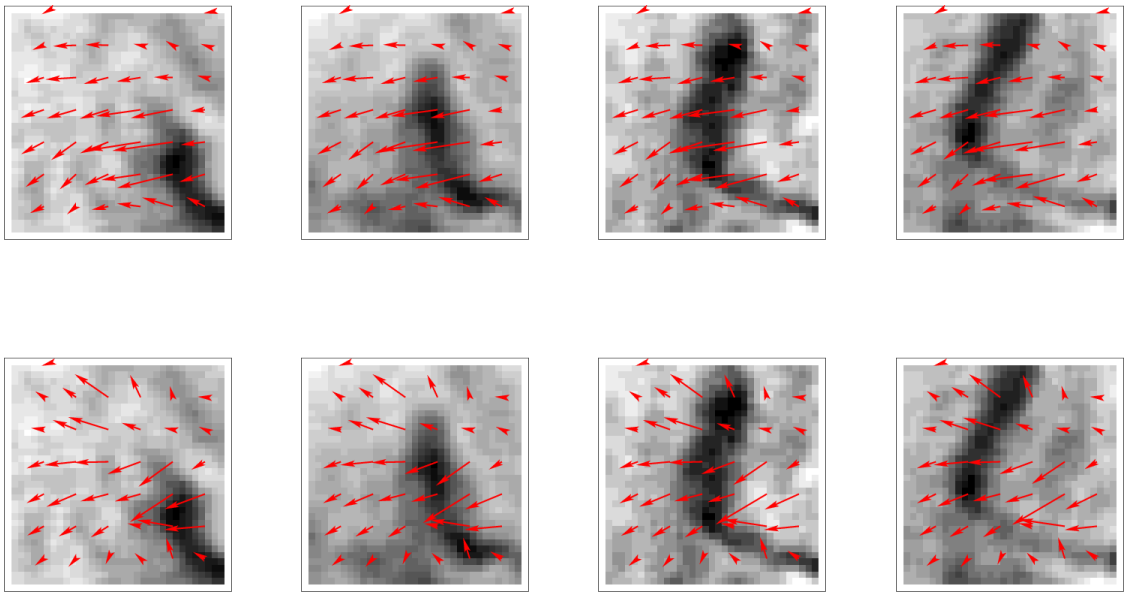

Figure 5. Top: Non-linear observation equation. Bottom: linear.

- Data assimilation in the full model is applied on the first window to retrieve the motion field.
- The estimate of motion field at the beginning of the second window makes it possible to define the subspace for motion and a reduced model is obtained by Galerkin projection.
- Data assimilation in the reduced model is applied for this second window.
- The process is then iterated for the next window until the end of the whole image sequence.

Experiments were designed to quantify the results of this sliding windows method with base obtained by Principal Orthogonal Decomposition or computed as bi-sine functions [29].

6.6.5. 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 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.

Part of research was concerned on the numerical schemes applied for advecting the distance map and designing its adjoint.

6.7. Minimax filtering

In minimax filtering for state estimation, the initial state error, the model error and the observational errors are classically supposed to belong to one joint ellipsoid. In this case, it is only assumed that the errors, stochastic or deterministic, are bounded. For each assimilation experiment, the filter computes an ellipsoid where one will find at least all states compatible with observations and errors description. The state estimate is taken as the center of the ellipsoid. No assumption on the actual distribution of the errors in needed and the state estimate minimizes the worst-case error, which makes the filter robust.

6.7.1. A posteriori minimax motion estimation

Participants: Sergiy Zhuk [IBM Research, Ireland], Isabelle Herlin, Olexander Nakonechnyi [Taras Shevchenko National University of Kyiv], Jason Frank [CWI, the Netherlands].

Data assimilation algorithms based on the 4D-Var formulation look for the so-called conditional mode estimate. The latter maximizes the conditional probability density function, provided the initial condition, model error and observation noise are realizations of independent Gaussian random variables. However this Gaussian assumption is often not satisfied for geophysical flows. Moreover, the estimation error of the conditional mode estimate is not a first-hand result of these methods. The issues above can be addressed by means of the Minimax State Estimation (MSE) approach. It allows to filter out any random (with bounded correlation operator) or deterministic (with bounded energy) noise and assess the worst-case estimation error.

The iterative MSE algorithm was developed for the problem of optical flow estimation from a sequence of 2D images. The main idea of the algorithm is to use the "bi-linear" structure of the Navier-Stokes equations and optical flow constraint in order to iteratively estimate the velocity. The algorithm consists of the following parts:

1) we construct pseudo-observations $\widehat{I}$ as the estimate of the image brightness function $I(x, y, t)$ solving the optical flow constraint such that $\widehat{I}$ fits (in the sense of least-squares) the observed sequence of images. To do so, we set the velocity field in the optical flow constraint to be the current minimax estimate of the velocity field $\mathbf{w}$, obtained at the previous iteration of the algorithm, and construct the minimax estimate $\widehat{I}$ of the solution of the resulting linear advection equation using the observed image sequence as discrete measurements of the brightness function;

2) we plug the estimate of the image gradient, obtained out of pseudo-observations $\widehat{I}$ in 1), into the optical flow constraint and the current minimax estimate $\mathbf{w}$ of the velocity field into the non linear part of Navier-Stokes equations so that we end up with a system of linear PDEs, which represents an extended state equation: it contains a linear parabolic equation for the velocity field and linear advection equation for the image brightness function; we construct the minimax estimate of the velocity field from the extended state equation using again the observed image sequence as discrete measurements of the brightness function;

3) we use the minimax estimate of the velocity field obtained in 2) in order to start 1) again.

Point 1) has been implemented and tested. As Point 2) is currently under development, it is replaced by one of our motion estimation method in order to be plugged in Point 3).

6.8. Fire application

6.8.1. Model evaluation for fire propagation

Participants: Vivien Mallet, Jean-Baptiste Fillipi [CNRS], Bahaa Nader [University of Corsica].

In the field of forest fires risk management, important challenges exist in terms of people and goods preservation. Answering to strong needs from different actors (firefighters, foresters), researchers focus their efforts to develop operational decision support system tools that may forecast wildfire behavior. This requires the evaluation of models performance, but currently, simulation errors are not sufficiently qualified and quantified. As the main objective is to realize a *decision support system*, it is required to establish robust forecast evaluations. In the context of the ANR project IDEA, the evaluation of model simulations has led to the definition and implementation of a series of forecast scores. The merits and shortcomings of the scores were evaluated on synthetic cases. This demonstrated the efficiency of scores that take into account the fire dynamics, where some classical scores may fail. This was also found on real fires, using field observations.

In addition, we consider that the proper evaluation of a model requires to apply it to a large number of fires – instead of carrying out a fine tuning on just one fire. We implemented a software to simulate a large number of fires (from the Prométhée database, http://www.promethee.com/) with the simulation model ForeFire (CNRS/University of Corsica) and evaluate the results with error measures. One simulation requires mainly the following data: the ignition point, the ground elevation, the vegetation cover and the wind field. See illustration on Fig. 6 .

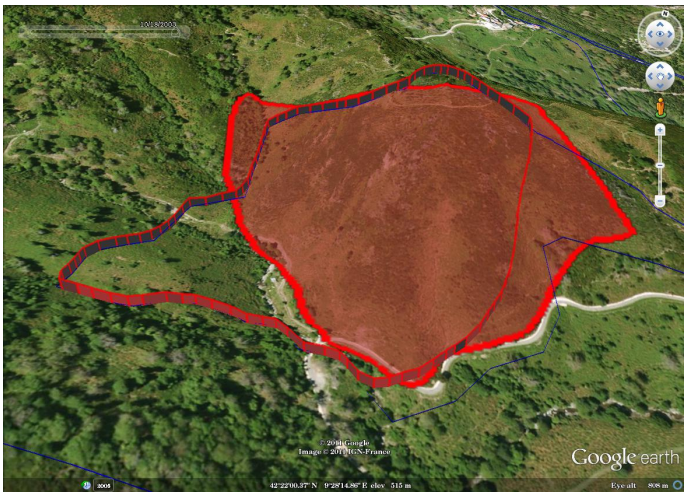

Figure 6. Fire simulation (using ForeFire) in red elevated contour, and observation (from Prométhée) of the burned area in filled red contour, for a 2003 fire near San-Giovanni-di-Moriani (Corsica).

POMDAPI Project-Team (section vide)

REO Project-Team

6. New Results

6.1. Mathematical and numerical analysis of fluid-structure interaction problems

Participants: Cristóbal Bertoglio Beltran, Muriel Boulakia, Miguel Ángel Fernández Varela, Sébastien Martin, Jean-Frédéric Gerbeau, Jimmy Mullaert, Marina Vidrascu.

- In [26], we study a three-dimensional fluid-structure interaction problem. The motion of the fluid is modeled by the Navier-Stokes equations and we consider for the elastic structure a finite-dimensional approximation of the equation of linear elasticity. The time variation of the fluid domain is not known a priori, so we deal with a free boundary value problem. Our main result yields the local in time existence and uniqueness of strong solutions for this system.
- In [31], a robust finite volume method for the solution of high-speed compressible flows in multi-material domains involving arbitrary equations of state and large density jumps is presented. One of the main contributions of this paper is a tabulation method based on a sparsegrid approximation to solve very efficiently two-phase Riemann problems for arbitrary equations of state. The proposed computational method is illustrated with the three-dimensional simulation of the dynamics of an underwater explosion bubble.
- In [52] we analyze the performances of several Luenberger observers to estimate the state of a fluid-structure interaction model for hemodynamics, when the measurements are assumed to be restricted to displacements or velocities in the solid. The present framework establishes that these methods are very attractive strategies (compared, e.g., to classical variational techniques) to perform state estimation.
- In [51] we analyze two 3D-0D coupling approaches in which a fractional-step projection scheme is used in the fluid. We introduce and analyze an implicitly 3D-0D coupled formulation with enhanced stability properties and which requires a negligible additional computational cost. The theoretical stability results are confirmed by meaningful numerical experiments in patient specific geometries coming from medical imaging.
- In [55] we introduce a class of explicit Robin-Neumann schemes for the explicit coupling of a general thin-structure (e.g., viscoelastic and non-linear) with an incompressible fluid. These methods generalize the displacement correction schemes introduced in [32]. A priori stability and convergence error estimates show that optimal first-order accuracy can be achieved with appropriate extrapolation and without compromising stability. A deep numerical study confirms the theoretical findings.
- In [64] we present two-dimensional simulations of chemotactic self-propelled bacteria swimming in a viscous fluid. Self-propulsion is modelled by a couple of forces of same intensity and opposite direction applied on the rigid bacterial body and on an associated region in the fluid representing the flagellar bundle. The orientations of the individual bacteria are subjected to random changes, with a frequency that depends on the surrounding oxygen concentration, in order to favor the direction of the concentration gradient.
- In [40] we propose a method of modeling sail structures which captures the wrinkling behavior of such structures. The method is validated through experimental and analytical test cases, particularly in terms of wrinkling prediction. An enhanced wrinkling index is proposed as a valuable measure characterizing the global wrinkling development on the deformed structure. The method is based on a pseudo-dynamic finite element procedure involving non-linear MITC shell elements. The major advantage compared to membrane models generally used for this type of analysis is that no ad hoc

wrinkling model is required to control the stability of the structure. We demonstrate our approach to analyse the behavior of various structures with spherical and cylindrical shapes, characteristic of downwind sails over a rather wide range of shape constitutive parameters. In all cases convergence is reached and the overall flying shape is most adequately represented, which shows that our approach is a most valuable alternative to standard techniques to provide deeper insight into the physical behaviour. Limitations appear only in some very special instances in which local wrinkling-related instabilities are extremely high and would require specific additional treatments.

6.2. Numerical methods for fluid mechanics and application to blood flows

Participants: Grégory Arbia, Jean-Frédéric Gerbeau, Sébastien Martin, Saverio Smaldone, Marc Thiriet, Irène Vignon-Clementel.

- In [18], a procedure for modeling the heart valves is presented. Instead of modeling complete leaflet motion, leaflets are modeled in open and closed configurations. This method enables significant computational savings compared to complete fluid-structure interaction and contact modeling, while maintaining realistic three-dimensional velocity and pressure distributions near the valve, which is not possible from lumped parameter modeling. To illustrate the versatility of the model, realistic and patient-specific simulations are presented, as well as comparison with complete fluid-structure interaction simulation.
- [37] paves the way for a complete patient-specific fluid-structure vascular modeling in which all types of available measurements could be used to estimate uncertain parameters of biophysical and clinical relevance. We propose a complete methodological chain for the identification of the parameters involved in a model for external tissue support of blood vessels, using patient image data. We demonstrate the use of this framework in a realistic application case involving hemodynamics in the thoracic aorta. The estimation of the boundary support parameters proves successful, in particular in that direct modeling simulations based on the estimated parameters are more accurate than with a previous manual expert calibration.
- In [27] we study the image-based blood flow in the first generation of the pulmonary arterial tree. This patient-specific study is aimed at assessing effects of lung deformation and vascular resistance on the pulmonary blood flow, especially during the acute phase of a pneumothorax and after recovery. Arterial geometry was extracted up to the fifth generation from computed tomography images, and reconstructed. An unsteady laminar flow with a given set of resistances at outlets was modeled. Adaptation is set to match perfusion to ventilation.
- In [44], [36] we study the reciprocal effect of blood circulation and high-intensity focused ultrasound on the temperature field in the liver. High-intensity focused ultrasound (HIFU) is used as a thermal ablation process to eliminate tumors in different body's organs. Blood flow has a cooling effect. Conversely, ultrasounds are responsible for acoustic streaming. A three-dimensional acoustics-thermal-fluid coupling model is carried out to compute the temperature field a given hepatic cancerous region.
- The use of elaborate closed-loop lumped parameter network (LPN) models of the heart and the circulatory system as boundary conditions for 3D simulations can provide valuable global dynamic information, particularly for patient specific simulations. In [30], we have developed and tested a numerical method to couple a 3D Navier-Stokes finite-element formulation and a reduced model of the rest of the circulation, keeping the coupling robust but modular. For Neumann boundaries, implicit, semi-implicit, and explicit quasi-Newton formulations are compared within the time-implicit coupling scheme. The requirements for coupling Dirichlet boundary conditions are also discussed and compared to that of the Neumann coupled boundaries. Both these works were key for applications where blood flows in different directions during the cardiac cycle and where coupling with the rest of the circulation is instrumental (see the shunt optimization application [29]).

- Boundary conditions in patient-specific blood flow simulations is key because pressure and flow within the modeled domain are driven by the interplay between the local 3D hemodynamics and the rest of the circulation. However, these boundary conditions are rarely the measured variables. In [45], we showed how one can go from patient-specific clinical data (MRI and catheterization) to simulation input parameters, including modeling assumptions and the impact of both on simulation results. We explained how Windkessel models and more involved LPN can be calibrated.

  In [34], we developed two multi-scale models, each including the 3D model of the surgical junction constructed from MRI, and a closed-loop LPN derived from pre-operative data obtained from two patients prior to Stage 2 Fontan palliation of single ventricle congenital heart disease. "Virtual" surgeries were performed and a corresponding multi-scale simulation predicted the patient's post-operative hemodynamic conditions, tested under different physiological conditions. The impact of the surgical junction geometry on the global circulation was contrasted with variations of key physiological parameters.
- In [19], a similar 3D multiscale model was used but for the Stage 3 Fontan palliation. Several studies have been done to optimize the geometry of the surgical connection, to minimizing energy losses and improving surgical outcomes, but usually without taking into account respiration or exercise. A respiration model that modulates the extravascular pressures in the thoracic and abdominal cavities was implemented. Results showed that the preoperative model is able to realistically capture cardiac and respiratory oscillations compared to the venous Doppler velocity tracings. Three virtual surgical alternatives were coupled to the LPN and then investigated under rest and exercise conditions.
- In [29], such a 3D-closed loop LPN model was integrated with an automated derivative-free optimization algorithm in an idealized systemic-to-pulmonary shunt anatomy (Stage 1 Fontan palliation). The goal was to optimize shunt geometries. Clinicians selected three objective functions to be maximized: (1) systemic, (2) coronary, and (3) combined systemic and coronary oxygen. Results showed the geometries associated with the favored delivery, the origin of coronary artery flow being driven by the shunt position as well. The results made only sense when the 3D domain was connected to a closed-loop model of the circulation.
- A novel Y-shaped baffle was proposed for the Stage 3 Fontan operation achieving overall superior hemodynamic performance compared with traditional designs. Previously, we investigated if and how the inferior vena cava flow (which contains an important biological hepatic factor) could be best distributed among both lungs. In [41] we proposed a multi-step method for patient-specific optimization of such surgeries to study the effects of boundary conditions and geometry on hepatic factor distribution (HFD). The resulting optimal Y-graft geometry largely depended on the patient left/right pulmonary flow split. Unequal branch size and constrained optimization on energy efficiency were explored. Two patient-specific examples showed that optimization-derived Y-grafts effectively improved HFD.

6.3. Numerical methods for cardiac electrophysiology

Participants: Muriel Boulakia, Miguel Ángel Fernández Varela, Jean-Frédéric Gerbeau, Vincent Martin, Elisa Schenone.

- In [62], we propose a surface-based electrophysiology formulation, motivated by the modeling of thin structures such as cardiac atria, which greatly reduces the size of the computational models. Our model is specifically devised to retain the key features associated with the anisotropy in the diffusion effects induced by the fiber architecture, with rapid variations across the thickness which cannot be adequately represented by naive averaging strategies. The model relies on a detailed asymptotic analysis in which we identify a limit model and establish strong convergence results. We also provide detailed numerical assessments which confirm an excellent accuracy of the surface-based model – compared with the reference 3D model – including in the representation of a complex phenomenon, namely, spiral waves.

6.4. Lung and respiration modeling

Participants: Laurent Boudin, Paul Cazeaux, Bérénice Grec, Muriel Boulakia, Anne-Claire Egloffe, Benoit Fabreges, Miguel Ángel Fernández Varela, Jean-Frédéric Gerbeau, Céline Grandmont, Stéphane Liwarek, Sébastien Martin, Ayman Moussa.

- [59], [60]:We are concerned here with identifiability, stability properties and estimates for the inverse problem of identifying a Robin coefficient on some non accessible part of the boundary from available data on the other part of boundary corresponding to solutions of the Stokes equations. In [59], we first consider a steady state two-dimensional Stokes problem and study the identifiability of Robin coefficient and then we establish a stability estimate of logarithm type using a global Carleman inequality. We then consider the unsteady problem. In [60]:We prove hölderian and logarithmic stability estimates associated to the unique continuation property for the Stokes system. The proof of these results is based on local Carleman inequalities. In the second part, these estimates on the fluid velocity and on the fluid pressure are applied to solve the inverse problem of identifying a Robin coefficient. For this identification parameter problem, we obtain a logarithmic stability estimate under the assumption that the velocity of a given reference solution stays far from 0 on a part of the boundary where Robin conditions are prescribed.
- In [61] we are interested in the mathematical modeling of the propagation of sound waves in the lung parenchyma, which is a foam–like elastic material containing millions of air– filled alveoli. In this study, the parenchyma is governed by the linearized elasticity equations and the air by the acoustic wave equations. The geometric arrangement of the alveoli is assumed to be periodic with a small period $\varepsilon > 0$. We consider the time–harmonic regime forced by vibrations induced by volumic forces. We use the two–scale convergence theory to study the asymptotic behavior as ε goes to zero and prove the convergence of the solutions of the coupled fluid–structure problem to the solution of a linear–elasticity boundary value problem.
- In [53] we develop and study numerically a model to describe some aspects of sound propagation in the human lung, considered as a deformable and viscoelastic porous medium (the parenchyma) with millions of alveoli filled with air. Transmission of sound through the lung above 1 kHz is known to be highly frequency–dependent. We pursue the key idea that the viscoelastic parenchyma structure is highly heterogeneous on the small scale ε and use two–scale homogenization techniques to derive effective acoustic equations for asymptotically small ε. This process turns out to introduce new memory effects. The effective material parameters are determined from the solution of frequency–dependent micro–structure cell problems. We propose a numerical approach to investigate the sound propagation in the homogenized parenchyma using a Discontinuous Galerkin formulation. Numerical examples are presented.
- In [22], we consider the Maxwell-Stefan model of diffusion previously introduced. We provide a qualitative and quantitative mathematical and basic numerical analysis of the model.
- In [65] we propose an integrated model for oxygen transfer into the blood, coupled with a lumped mechanical model for the ventilation process. We aim at investigating oxygen transfer into the blood at rest or exercise. The first task consists in describing nonlinear effects of the oxygen transfer under normal conditions. We also include the possible diffusion limitation in oxygen transfer observed in extreme regimes involving parameters such as alveolar and venous blood oxygen partial pressures, capillary volume, diffusing capacity of the membrane, oxygen binding by hemoglobin and transit time of the red blood cells in the capillaries. The second task consists in discussing the oxygen concentration heterogeneity along the path length in the acinus.
- In [43] we presented preliminary work on a multiscale 3D-0D airflow model to study differences between healthy and emphysema rats. The 0D model parameters were estimated from experimental data. 3D Navier-Stokes simulations were performed in healthy lungs, and in homogenous and heterogeneous emphysema lungs.

6.5. Miscellaneous

Participants: Laurent Boudin, Jean-Frédéric Gerbeau, Damiano Lombardi, Sébastien Martin, Marina Vidrascu, Irène Vignon-Clementel.

- In [56], a reduced-order model algorithm, based on approximations of Lax pairs, is proposed to solve nonlinear evolution partial differential equations. Contrary to other reduced-order methods, like Proper Orthogonal Decomposition, the space where the solution is searched for evolves according to a dynamics specific to the problem. It is therefore well-suited to solving problems with progressive waves or front propagation. Numerical examples are shown for the KdV and FKPP (nonlinear reaction diffusion) equations, in one and two dimensions.
- In [21], we investigate the asymptotic behaviour of the solutions to the non-reactive fully elastic Boltzmann equations for mixtures in the diffusive scaling. We deal with cross sections such as hard spheres or cut-off power law potentials. We use Hilbert expansions near the common thermodynamic equilibrium granted by the H-theorem. The lower-order non trivial equality obtained from the Boltzmann equations leads to a linear functional equation in the velocity variable which is solved thanks to the Fredholm alternative. Since we consider multicomponent mixtures, the classical techniques introduced by Grad cannot be applied, and we propose a new method to treat the terms involving particles with different masses. The next-order equality in the Hilbert expansion then allows to write the macroscopic continuity equations for each component of the mixture.
- In [58], we discuss some numerical properties of the viscous numerical scheme introduced in [23] to solve the one-dimensional pressureless gases system, and study in particular, from a computational viewpoint, its asymptotic behavior when the viscosity parameter $\varepsilon > 0$ used in the scheme becomes smaller.
- In [33] we study a network-based model for rubber. Since the pioneering work by Treloar, many models based on polymer chain statistics have been proposed to describe rubber elasticity. Recently, Alicandro, Cicalese, and the first author rigorously derived a continuum theory of rubber elasticity from a discrete model by variational convergence. The aim of this paper is twofold. First we further physically motivate this model, and complete the analysis by numerical simulations. Second, in order to compare this model to the literature, we present in a common language two other representative types of models, specify their underlying assumptions, check their mathematical properties, and compare them to Treloar's experiments.
- In [63] our aim is to demonstrate the effectiveness of the matched asymptotic expansion method in obtaining a simplified model for the influence of small identical heterogeneities periodically distributed on an internal surface on the overall response of a linearly elastic body. The results of some numerical experiments corroborate the precise identification of the different steps, in particular of the outer/inner regions with their normalized coordinate systems and the scale separation, leading to the model.
- In cancer modeling, to be able to capture the full in-vivo scale, tumors have to be modeled with continuum models. An important step consists in qualitatively and quantitatively comparing agent-based models (which parameters can generally be identified by experiments in vitro) and continuum models. We derived a first 1D continuum model for tumor growth from the cell based model (Drasdo and Hoehme, 2005): it results in a fluid-type model which capture tumor expansion in both diffusive and compact phenotypes. The tumor expands based on the pressure gradient generated by cell proliferation, the latter being hindered by high density or pressure. In [39] this modeled is analyzed mathematically, showing the existence of traveling waves in the different regimes (with or without internal friction and diffusion due to active movement). In particular the incompressible cells limit is very singular and relates to the Hele-Shaw equation. Numerical results confirm the analysis.

SISYPHE Project-Team

5. New Results

5.1. Modeling, observation and control: systems modeled by ordinary differential equations

5.1.1. Nonlinear system identification

Participants: Pierre-Alexandre Bliman, Boyi Ni, Michel Sorine, Qinghua Zhang.

In the framework of the joint Franco-Chinese ANR-NSFC EBONSI project, in collaboration with the Laboratory of Industrial Process Monitoring and Optimization of Peking University, and with Centre de Recherche en Automatique de Nancy (CRAN), the topics studied this year on nonlinear system identification are mainly on extended Hammerstein system identification with hysteresis nonlinearity and on continuous time block-oriented nonlinear system identification.

Motivated by the modeling of control valves with significant stiction, we have studied extended Hammerstein systems composed of a hysteresis nonlinearity followed by a linear dynamic subsystem. The joint characterization of the control valve and of the controlled process is formulated as the identification of an extended Hammerstein system. A point-slope based hysteresis model is used to describe the input hysteresis nonlinearity of the control valve. An iterative algorithm is proposed to solve the identification problem. The basic idea is to separate the ascent and descent paths of the input hysteresis nonlinearity subject to oscillatory excitations. Industrial examples are tested to verify the effectiveness of the proposed identification algorithm for characterizing complex behavior of control valve stiction in practice. This work has been presented at the 16th IFAC Symposium on System Identification [66].

A Hammerstein-Wiener system is composed of a dynamic linear subsystem preceded and followed by two static nonlinearities. Typically, the nonlinearities of such a system are caused by actuator and sensor distortions. The identification of such systems with a continuous time model has been studied this year in collaboration with colleagues of CRAN. Based on previously developed simplified refined instrumental variable method, and by making use of an adaptive observer for data filtering, a combined approach, referred to as Kalman pre-filtered instrumental variable based method, is developed. By taking advantages of the two aforementioned methods, the new method is faster and has a naturally stabilizing Kalman filter that does not color white noises. This work has been presented at the 16th IFAC Symposium on System Identification [62].

5.1.2. Model-based fault diagnosis

Participants: Abdouramane Moussa Ali, Qinghua Zhang.

The increasing requirements for higher performance, efficiency, reliability and safety of modern engineering systems call for continuous research investigations in the field of fault detection and isolation. This year we have studied algebro-differential systems through an adaptive observer based approach, and linear time varying systems through a Kalman filter based statistical testing approach.

In the framework of the MODIPRO project funded by Paris Region ASTech, the monitoring of the air conditioning system of an aircraft has been studied this year. Part of this system is modeled by nonlinear algebro-differential equations. A method for fault diagnosis of such systems has been developed in our study. Through a particular discretization method and under realistic assumptions, the considered continuous time DAE model is transformed to an explicit state space model in discrete time. An adaptive observer is then applied to the discretized system for monitoring faults possibly affecting the system and represented by changes in model parameters. This work will be presented at the 5th IFAC Symposium on System Structure and Control [61].

While the theory of fault diagnosis has been mostly developed for linear time invariant (LTI) systems, in many industrial applications it is important to take into account the nonlinear behavior of the monitored systems. One possible approach is to linearize a nonlinear system all along its state trajectory, resulting in linear time varying (LTV) or linear parameter varying (LPV) models. In collaboration with Michèle Basseville of IRISA (Institut de Recherche en Informatique et Systèmes Aléatoires), fault diagnosis for stochastic LTV systems has been studied this year. By applying the Kalman filter in a particular manner avoiding the difficulty related to unknown faults possibly affecting the system, the problem of fault diagnosis in a dynamic LTV system is transformed into a hypothesis testing problem in a simple linear regression model. Generalized likelihood ratio (GLR) tests are then applied to the resulting hypothesis testing problem. This work has been presented at the 16th IFAC Symposium on System Identification [67].

5.2. Observation, control and traveling waves in systems modeled by partial differential equations

5.2.1. Modeling of electric transmission networks

Participants: Mohamed Oumri, Michel Sorine, Qinghua Zhang.

The increasing number and complexity of wired electric networks in modern engineering systems is amplifying the importance of the reliability of electric connections. In the framework of the ANR 0-DEFECT project, we have studied mathematical models of complex electric networks with the aim of designing algorithms for fault diagnosis. The well known Baum-Liu-Tesche (BLT) equation is a powerful model for describing quite general networks and allows to compute the current and voltage waves at the nodes of a network from the specifications of its nodes and connecting cables [63]. This year we have studied the inverse problem: what can we know about the properties of the cables connecting the nodes of a network from experiments made at the nodes of the network? A convenient model for this purpose is formulated with admittance matrices. It is essentially equivalent to the BLT equation, hence can describe quite general networks. The inverse problem is then solved through a decomposition of the admittance matrix of the entire network.

5.2.2. Diagnosis of insulator degradation in long electric cables

Participants: Leila Djaziri, Michel Sorine, Qinghua Zhang.

For the diagnosis of insulator degradation in long electric cables, the estimation of the shunt conductance of such cables have been studied, in the framework of the ANR INSCAN project. The shunt conductance of a healthy electric cable is usually very weak. Even when the insulator in the cable is significantly degraded, the shunt conductance can still remain at a quite low level. The main difficulty in this study is due to the fact that the measurements made at the ends of a cable are hardly sensitive to the variations of the shunt conductance. To overcome this difficulty, two methods have been studied. One of them is based on the analysis of the sensitivity of the wave phase shift to the shunt conductance. The efficiency of this method has been demonstrated through extensive tests on cables of SNCF (Société Nationale des Chemins de Fer français). Another method is based on the processing of long time data records. It is designed for the estimation of distributed shunt conductance, in order to detect and to locate inhomogeneous degradation of the insulator. The main idea of this method is to compensate the weak sensitivity of the measurement by long time data records. The results of this method evaluated by numerical simulations have been reported at the 16th IFAC Symposium on System Identification [68].

5.3. System theory approach of some quantum systems

Participants: Hadis Amini, Zaki Leghtas, Mazyar Mirrahimi, Pierre Rouchon.

Most of this work is done in close collaboration with the Pierre Aigrain laboratory (LPA) at ENS Paris and the Quantronics Laboratory (Qlab) of Michel Devoret and the Rob Schoelkopf Lab at Yale University.

Modern scientific and technologic requirements have led the theoretical and experimental research toward an engineering of quantum systems. The technologies that are proposed or developed include nano-scale electromechanical devices, tools for implementing quantum computation and quantum communication, NMR applications, quantum chemistry synthesis, high-resolution sensors, etc. The recent theoretical and experimental researches have shown that the quantum dynamics can be studied in the framework of the theory of estimation and control of systems, but give place to models that are not completely explored yet.

Our activities lie in the theoretical and experimental interface of this progressing field of research with an accent on the applications in quantum information and computation as well as high-precision metrology. By focusing on two different but similar types of experimental setups, consisting of cavity quantum electrodynamical systems and quantum Josephson circuits, we aim in preparing highly non-classical states of a microwave field and protect these states against decoherence. Two different approaches are considered: 1- real-time measurement, quantum filtering and feedback ; 2- dissipation engineering also called reservoir engineering. Through the first methodology, we try to propose new experimental feedback protocols based on a fast real-time processing of measurement signal, followed by a state estimation applying the filtered signal and finally designing simple feedback laws based on the estimated state. The second methodology consists of designing new quantum circuit schemes that allow to orient the system's coupling to its environment in such a way that evacuates the undesired entropy induced by un-controlled noise sources.

5.3.1. Measurement based feedback

We have developed the mathematical methods underlying a recent quantum feedback experiment stabilizing photon-number states [17], [30], [29], [24]. We consider a controlled system whose quantum state, a finite dimensional density operator, is governed by a discrete-time nonlinear Markov process. In open-loop, the measurements are assumed to be quantum non-demolition (QND) measurements. This Markov process admits a set of stationary pure states associated to an orthonormal basis. These stationary states provide martingales crucial to prove the open-loop stability: under simple assumptions, almost all trajectories converge to one of these stationary states; the probability to converge to a stationary state is given by its overlap with the initial quantum state. From these open-loop martingales, we construct a supermartingale whose parameters are given by inverting a Metzler matrix characterizing the impact of the control input on the Kraus operators defining the Markov process. This supermartingale measures the "distance" between the current quantum state and the goal state chosen from one of the open-loop stationary pure states. At each step, the control input minimizes the conditional expectation of this distance. It is proven that the resulting feedback scheme stabilizes almost surely towards the goal state whatever the initial quantum state. This state feedback takes into account a known constant delay of arbitrary length in the control loop. This control strategy is proved to remain also convergent when the state is replaced by its estimate based on a quantum filter. It relies on measurements that can be corrupted by random errors with conditional probabilities described by a known left stochastic matrix. Closed-loop simulations corroborated by experimental data illustrate the interest of such nonlinear feedback scheme for the photon box [29].

We have also investigated the stabilization of the dynamical state of a superconducting qubit. In a series of papers, A. Korotkov and his co-workers suggested that continuous weak measurement of the state of a qubit and applying an appropriate feedback on the amplitude of a Rabi drive, should maintain the coherence of the Rabi oscillations for arbitrary time. Here, in the aim of addressing a metrological application of these persistent Rabi oscillations, we explore a new variant of such strategies. This variant is based on performing strong measurements in a discrete manner and using the measurement record to correct the phase of the Rabi oscillations. Noting that such persistent Rabi oscillations can be viewed as an amplitude- to-frequency convertor (converting the amplitude of the Rabi microwave drive to a precise frequency), we propose another feedback layer consisting of a simple analog phase locked loop to compensate the low frequency deviations in the amplitude of the Rabi drive [60].

5.3.2. Dissipation engineering

We have introduced a new quantum gate that transfers an arbitrary state of a qubit into a superposition of two quasi-orthogonal coherent states of a cavity mode, with opposite phases. This qcMAP gate is based on

conditional qubit and cavity operations exploiting the energy level dispersive shifts, in the regime where they are much stronger than the cavity and qubit linewidths [77], [26]. The generation of multi-component superpositions of quasi-orthogonal coherent states, non-local entangled states of two resonators and multi-qubit GHZ states can be efficiently achieved by this gate. We also propose a new method, based on the application of this gate, to autonomously correct for errors of a logical qubit induced by energy relaxation. This scheme encodes the logical qubit as a multi-component superposition of coherent states in a harmonic oscillator. The error correction is performed by transferring the entropy to an ancila qubit and reseting the qubit. We layout in detail how to implement these operations in a practical system [78]. This proposal directly addresses the task of building a hardware-efficient and technically realizable quantum memory [78].

We have also studied the application of dissipation engineering techniques to perform a high-performance and fast qubit reset. Qubit rest is crucial at the start of and during quantum information algorithms. Our protocol, nicknamed DDROP (Double Drive Reset of Population) is experimentally tested on a superconducting transmon qubit and achieves a ground state preparation of at least 99.5% in times less than 3μs; faster and higher fidelity are predicted upon parameter optimization [74]. We are currently working on extending our protocol to prepare and protect two-qubit entangled states and to perform autonomous quantum error correction.

5.4. Modeling, observation and control in biosciences - Reproductive system

5.4.1. Numerical simulation of the selection process of the ovarian follicles

Participants: Benjamin Aymard, Frédérique Clément.

Collaboration with Frédéric Coquel and Marie Postel.

Implementation of a parallelized numerical scheme based on finite volumes. We have designed and implemented a numerical method to simulate a multiscale model describing the selection process in ovarian follicles [11], [10]. The PDE model consists in a quasi-linear hyperbolic system of large size, namely $N_f \times N_f$, ruling the time evolution of the cell density functions of N_f follicles (in practice N_f is of the order of a few to twenty). These equations are weakly coupled through the sum of the first order moments of the density functions. The time-dependent equations make use of two structuring variables, age and maturity, which play the roles of space variables. The problem is naturally set over a compact domain of $\mathbf{R}^2$. The formulation of the time-dependent controlled transport coefficients accounts for available biological knowledge on follicular cell kinetics. We introduce a dedicated numerical scheme that is amenable to parallelization, by taking advantage of the weak coupling. Numerical illustrations assess the relevance of the proposed method both in term of accuracy and HPC achievements [32].

A numerical method for cell dynamics: kinetic equations with discontinuous coefficients. The motivation of this work is the numerical treatment of the mitosis in biological models involving cell dynamics. More generally we study hyperbolic PDEs with flux transmission conditions at interfaces between subdomains where coefficients are discontinuous. A dedicated finite volume scheme with a limited high order enhancement is adapted to treat the discontinuities arising at interfaces. The validation of the method is done on 1D and 2D toy problems for which exact solutions are available, allowing us to do a thorough convergence study. A simulation on the original biological model illustrates the full potentialities of the scheme [72].

5.4.2. Optimal control of cell mass and maturity in a model of follicular ovulation

Participants: Frédérique Clément, Peipei Shang.

Collaboration with Jean-Michel Coron

We have studied some optimal control problems associated with a scalar hyperbolic conservation law modeling the development of ovarian follicles. Changes in the age and maturity of follicular cells are described by a 2D conservation law, where the control terms act on the velocities. The control problem consists in optimizing the follicular cell resources so that the follicular maturity reaches a maximal value in fixed time. Formulating the optimal control problem within a hybrid framework, we have proved necessary optimality conditions in the form of Hybrid Maximum Principle [36]. We have then derived the optimal strategy and shown that there exists at least one optimal bang-bang control with one single switching time.

5.4.3. Multiscale analysis of mixed-mode oscillations in a phantom bursting model

Participants: Frédérique Clément, Mathieu Desroches, Maciej Krupa, Alexandre Vidal.

We have studied mixed mode oscillations in a model of secretion of GnRH (gonadotropin releasing hormone). The model is a phantom burster consisting of two feedforward coupled FitzHugh-Nagumo systems, with three time scales. The forcing system (Regulator) evolves on the slowest scale and acts by moving the slow null-cline of the forced system (Secretor). There are three modes of dynamics: pulsatility (transient relaxation oscillation), surge (quasi steady state) and small oscillations related to the passage of the slow null-cline through a fold point of the fast null-cline. We have derived a variety of reductions, taking advantage of the mentioned features of the system. We have obtained two results; one on the local dynamics near the fold in the parameter regime corresponding to the presence of small oscillations and the other on the global dynamics, more specifically on the existence of an attracting limit cycle. Our local result is a rigorous characterization of small canards and sectors of rotation in the case of folded node with an additional time scale, a feature allowing for a clear geometric argument. The global result gives the existence of an attracting unique limit cycle, which, in some parameter regimes, remains attracting and unique even during passages through a canard explosion [43].

5.4.4. A network model of the periodic synchronization process in the dynamics of calcium concentration in GnRH neurons

Participants: Frédérique Clément, Maciej Krupa, Alexandre Vidal.

Mathematical neuroendocrinology is a branch of mathematical neurosciences that is specifically interested in endocrine neurons, which have the uncommon ability of secreting neurohormones into the blood. One of the most striking features of neuroendocrine networks is their ability to exhibit very slow rhythms of neurosecretion, on the order of one or several hours. A prototypical instance is that of the pulsatile secretion pattern of GnRH (gonadotropin releasing hormone), the master hormone controlling the reproductive function, whose origin remains a puzzle issue since its discovery in the seventies. We have investigated the question of GnRH neuron synchronization on a mesoscopic scale and study how synchronized events in calcium dynamics can arise from the average electric activity of individual neurons. We have used as reference seminal experiments performed on embryonic GnRH neurons from rhesus monkeys, where calcium imaging series were recorded simultaneously in tens of neurons, and which have clearly shown the occurrence of synchronized calcium peaks associated with GnRH pulses, superposed on asynchronous, yet oscillatory individual background dynamics [100]. We have designed a network model by coupling 3D individual dynamics of FitzHugh-Nagumo type. Using phase-plane analysis, we have constrained the model behavior so that it meets qualitative and quantitative specifications derived from the experiments, including the precise control of the frequency of the synchronization episodes. In particular, we have shown how the time scales of the model can be tuned to fit the individual and synchronized time scales of the experiments. Finally, we have illustrated the ability of the model to reproduce additional experimental observations, such as partial recruitment of cells within the synchronization process or the occurrence of doublets of synchronization [76].

5.5. Clinical and physiological applications

5.5.1. DynPeak: An algorithm for pulse detection and frequency analysis in hormonal time series

Participants: Frédérique Clément, Claire Médigue, Alexandre Vidal, Qinghua Zhang.

Collaboration with Stéphane Fabre (UMR CNRS-INRA 6175).

The endocrine control of the reproductive function is often studied from the analysis of luteinizing hormone (LH) pulsatile secretion by the pituitary gland. Whereas measurements in the cavernous sinus cumulate anatomical and technical difficulties, LH levels can be easily assessed from jugular blood. However, plasma levels result from a convolution process due to clearance effects when LH enters the general circulation. Simultaneous measurements comparing LH levels in the cavernous sinus and jugular blood have revealed clear differences in the pulse shape, the amplitude and the baseline. Besides, experimental sampling occurs at a relatively low frequency (typically every 10 min) with respect to LH highest frequency release (one pulse per hour) and the resulting LH measurements are noised by both experimental and assay errors. As a result, the pattern of plasma LH may be not so clearly pulsatile. Yet, reliable information on the InterPulse Intervals (IPI) is a prerequisite to study precisely the steroid feedback exerted on the pituitary level. Hence, there is a real need for robust IPI detection algorithms. We have designed an algorithm for the monitoring of LH pulse frequency, basing ourselves both on the available endocrinological knowledge on LH pulse (shape and duration with respect to the frequency regime) and synthetic LH data generated by a simple model [54]. We make use of synthetic data to make clear some basic notions underlying our algorithmic choices. We focus on explaining how the process of sampling affects drastically the original pattern of secretion, and especially the amplitude of the detectable pulses. We then describe the algorithm in details and perform it on different sets of both synthetic and experimental LH time series. We further comment on how to diagnose possible outliers from the series of IPIs which is the main output of the algorithm.

ARLES Project-Team

6. New Results

6.1. Introduction

The ARLES project-team investigates solutions in the forms of languages, methods, tools and supporting middleware to assist the development of distributed software systems, with a special emphasis on mobile distributed systems enabling the ambient intelligence/pervasive computing vision. Our research activities in 2012 have focused on the following areas:

- Dynamic interoperability among networked systems toward making them eternal, by way of on-the-fly generation of connectors based on adequate system models (§ 6.2);
- Pervasive service-oriented software engineering, focusing on supporting service composition in an increasingly heterogeneous and dynamic networking environment, while enforcing quality of service (§ 6.3);
- Service oriented middleware for the ultra large scale future Internet of Things (§ 6.4);
- Abstractions for enabling domain experts to easily compose applications on the Internet of Things (§ 6.5); and
- The use of Requirement Engineering techniques for enabling systems to be self-adaptive under uncertainty (§ 6.6).

6.2. Emergent Middleware Supporting Interoperability in Extreme Distributed Systems

Participants: Emil Andriescu, Amel Bennaceur, Luca Cavallaro, Valérie Issarny, Daniel Sykes.

Interoperability is a fundamental challenge for today's extreme distributed systems. Indeed, the high-level of heterogeneity in both the application layer and the underlying infrastructure, together with the conflicting assumptions that each system makes about its execution environment hinder the successful interoperation of independently developed systems. A wide range of approaches have thus been proposed to address the interoperability challenge. However, solutions that require performing changes to the systems are usually not feasible since the systems to be integrated may be legacy systems, COTS (Commercial Off-The-Shelf) components or built by third parties; neither are the approaches that prune the behavior leading to mismatches since they also restrict the systems' functionality. Therefore, many solutions that aggregate the disparate systems in a non-intrusive way have been investigated. These solutions use intermediary software entities, called *mediators*, to interconnect systems despite disparities in their data and/or interaction models by performing the necessary coordination and translations while keeping them loosely-coupled. However, creating mediators requires a substantial development effort and a thorough knowledge of the application-domain, which is best understood by domain experts. Moreover, the increasing complexity of today's distributed systems, sometimes referred to as Systems of Systems, makes it almost impossible to develop 'correct' mediators manually. Therefore, formal approaches are used to synthesize mediators automatically.

In light of the above, we have introduced the notion of *emergent middleware* for realizing mediators. Our research on enabling emergent mediators is done in collaboration with our partners of the CONNECT project (§ 7.2.1.1). Our work during the year has more specifically focused on:

- **Architecture enabling emergent middleware.** We have been finalizing, together with our partners in the CONNECT project, the definition of an overall distributed system architecture supporting emergent middleware, from the discovery of networked systems to the learning of their respective behavior and synthesis of emergent middleware enabling them to interoperate [31].

- **Affordance inference.** We have proposed an ontology-based formal model of networked systems based on their affordances (high-level functionalities), interfaces, behavior, and non-functional properties, each of which describes a different facet of the system in a way similar to the service description promoted for semantic Web services. However, legacy systems do not necessarily specify all of the aforementioned facets. Therefore, we have explored techniques to infer the affordance by using textual descriptions of the interface of networked systems. More specifically, we rely on machine learning techniques to automate the inference of the affordance from the interface description by classifying the natural-language text according to a predefined ontology of affordances. In a complementary way, CONNECT partners investigate protocol-learning algorithms to learn the behavior of networked systems on the fly [17].
- **Mediator synthesis for emergent middleware.** We focus on systems that have compatible functionalities, i.e., semantically matching affordances, but are unable to interact successfully due to mismatching interfaces or behaviors. To solve such mismatches, we propose a *mapping-based* approach, whose goal is to automatically synthesize a mediator model that ensures the *safe* interaction of functionnally compatible systems, i.e., deadlock-freedom and the absence of unspecified receptions. Our approach combines semantic reasoning and constraint programming to identify the semantic correspondence between networked systems' interfaces, i.e., *interface mapping*. Unlike existing approaches that only tackle the one-to-one correspondence between actions and for which we investigated a solution using ontology-based model checking [16], the proposed mapping-based approach handles the more general cases of one-to-many and many-to-many mappings. This work has resulted in a supporting software prototype that allows validating the approach; related publication is under writing. A further key research issue we are addressing in emergent middleware is the study of cross-paradigm interaction so as to enable interoperability among highly heterogeneous services (e.g., an IT-based service will likely interact using the client-service scheme while thing-based services rather adopt asynchronous protocols). Toward that goal, we are studying abstract models associated with popular interaction paradigms and higher level, generic interaction paradigms to define cross-paradigm mappings that respect the behavioral semantics of the interacting systems.
- **Automated mediation for cross-layer protocol interoperability.** While existing approaches to interoperability consider either application or middleware heterogeneity separately, we believe that in real world scenarios this does not suffice: application and middleware boundaries are ill-defined and solutions to interoperability must consider them in conjunction. As part of our recent work, we have proposed such a solution, which solves cross-layer interoperability by automatically generating parsers and composers that abstract physical message encapsulation layers into logical protocol layers, thus supporting application layer mediation. Specifically, we support the automated synthesis of mediators at the application layer using the mapping-based approach discussed above, while we introduce *Composite Cross- Layer (CCL) parsers and composers* to handle cross-layer heterogeneity and to provide an abstract representation of the application data exchanged by the interoperating components. In particular, we associate the data embedded in messages with annotations that refers to concepts in a domain ontology. As a result, we are able to reason about the semantics of messages in terms of the operations and the data they require from or provide to one another and automatically synthesize, whenever possible, the appropriate mediators. We have demonstrated the validity of our approach by using the framework to solve cross-layer interoperability between existing conference management systems.
- **Models@run.time.** We have recently integrated the notion of *Models@run.time* in our research towards emergent middleware. We use Models@run.time to extend the applicability of models and abstractions to the runtime environment. As is the case for software development models, a run-time model is often created to support reasoning. However, in contrast to development models, run-time models are used to reason about the operating environment and runtime behavior, and thus these models must capture abstractions of runtime phenomena. Different dimensions need to be balanced, including resource-efficiency (time, memory, energy), context-dependency (time, location, platform), as well as personalization (quality-of-service specifications, profiles). The hypothesis is

that because Models@run.time provide meta-information for these dimensions during execution, run-time decisions can be facilitated and better automated.

Thus, we anticipate that Models@run.time will play an integral role in the management of extremely distributed systems. Our way of using runtime models captures syntax and also semantics of behaviour and supports runtime reasoning. Prior models@run.time approaches have generally concentrated on architectural-based runtime models and self-adaptation of existing software artifacts. However, such artefacts cannot always be produced in advance, and we believe that models@runtime have a fundamental role to play in the production of dynamic, adaptive, and on-the-fly software as investigated in the context of emergent middleware [8]. Specifically, two important methods underpin our approach: *i*) automatic inference of the required runtime models during execution and their refinement by exploiting learning and synthesis techniques; and *ii*) using these models for a dynamic software synthesis approach, where mediators are formally characterized (using LTS) to allow the runtime synthesis of software.

In order to enable emergent middleware, we have shown how systems can infer information to build runtime models during execution. Importantly, ontologies were exploited to enrich the runtime models and facilitated the mutual understanding required to perform the matching and mapping between the networked heterogeneous systems. Such reasoning about information that was not necessarily known before execution, is in contrast to the traditional use of models@run.time.

6.3. Revisiting the Abstractions of Service Oriented Computing for the Future Internet

Participants: Dionysis Athanasopoulos, Sandrine Beauche, George Bouloukakis, Oleg Davidyuk, Nikolaos Georgantas, Valérie Issarny, Ajay Kattepur.

A software architecture style characterizes, via a set of abstractions, the types of: components (i.e., units of computation or data stores), connectors (i.e., interaction protocols) and possibly configurations (i.e., system structures) that serve to build a given class of systems. As such, the definition of a software architectural style is central toward eliciting appropriate design, development and runtime support for any family of systems. The service oriented architecture style may then be briefly defined as follows: (1) components map to services, which may be refined into consumer, producer or prosumer services; (2) connectors map to traditional client-service interaction protocols; and (3) configurations map to compositions of services through (service-oriented) connectors, e.g., choreography and orchestration structures. While the service-oriented architecture style is well suited to support the development of Internet-based distributed systems, it is largely challenged by the Future Internet that poses new demands in terms of sustaining *ities* such as scalability, heterogeneity, mobility, awareness & adaptability that come in extreme degrees compared to the current Internet. Therefore, we have been working on eliciting software architectural abstractions for the Future Internet by building upon the service-oriented architecture style, as well as on applying them to system design, development and execution.

Complex distributed applications in the Future Internet will be to a large extent based on the open integration of extremely heterogeneous systems, such as lightweight embedded systems (e.g., sensors, actuators and networks of them), mobile systems (e.g., smartphone applications), and resource-rich IT systems (e.g., systems hosted on enterprise servers and Cloud infrastructures). These heterogeneous system domains differ significantly in terms of interaction paradigms, communication protocols, and data representation models, provided by supporting middleware platforms. Specifically considering interaction paradigms, the client/server (CS), publish/subscribe (PS), and tuple space (TS) paradigms are among the most widely employed ones today, with numerous related middleware platforms. In light of the above, we have aimed at eliciting abstractions that (i) leverage the diversity of interaction paradigms associated with today's and future complex distributed systems, as well as (ii) enable cross-paradigm interaction to sustain interoperability in the highly heterogeneous Future Internet.

Existing cross-domain interoperability efforts are based on bridging communication protocols, wrapping systems behind standard technology interfaces, and/or providing common API abstractions. In particular, such techniques have been applied by the two widely established system integration paradigms, that is, service oriented architecture (SOA) and enterprise service bus (ESB). However, state of the art interoperability efforts do not or only poorly address interaction paradigm interoperability. Indeed, systems integrated via SOA and ESB solutions have their interaction semantics transformed to the CS paradigm. Then, potential loss of interaction semantics can result in suboptimal or even problematic system integration. To overcome the limitation of today's ESB-based connectors for cross-domain interoperability in the Future Internet, we introduce a new connector type, called GA connector, which stands for "Generic Application connector". The proposed connector type is based on the service bus paradigm in that it achieves bridging across heterogeneous connector types. However, the behavior of the GA connector type differs from that of classical ESB connectors by bridging protocols across heterogeneous paradigms, which is further realized by paying special attention to the preservation of the semantics of the composed protocols. Indeed, the GA connector type is based on the abstraction and semantic-preserving merging of the common high-level semantics of base interaction paradigms.

Eliciting Interaction Paradigm Abstractions: We introduce a systematic abstraction of interaction paradigms with the following features:

- First, we introduce base CS, PS and TS connector types, which formally characterize today's core interaction paradigms. The proposed types comprehensively cover the essential semantics of the considered paradigms, based on a thorough survey of the related literature and representative middleware instances.
- Then, we further abstract these connector types into a single higher-level one, the GA connector type. GA is a comprehensive connector type based on the abstract union of CS, PS, and TS, where precise identification of the commonalities or similarities between the latter has enabled the optimization of the former. Further, GA preserves by construction the semantics of CS, PS, and TS.
- In more detail, connector types are formally specified in terms of: (i) their API (Application Programming Interface), and (ii) their roles, i.e., the semantics of interaction of the connected component(s) with the environment via the connector. Regarding the latter, the behavioral specification of roles from a middleware perspective relates to specifying the production and consumption of information in the network, while the semantics of the information are abstracted and dealt with at the application layer. The behaviors of the connector roles are then specified using Labeled Transition Systems (LTS). We precisely define the mapping of the roles implemented by the base connector types to/from the corresponding roles of the GA connector type.
- For both the above abstraction transformations, we provide counterpart concretizations, which enable transforming GA connector primitives to CS, PS, or TS connector primitives and then to concrete middleware platforms primitives.
- Furthermore, based on the GA abstraction, we introduce mapping transformations between any pair from the set {CS, PS, TS} via GA. The fine knowledge of CS, PS, and TS semantics, as embedded in GA, enables these mappings to be precise: differing semantics are mapped to each other in such a way that loss of semantics is limited to the minimum. These mappings relate to the definition of the glue process implemented by the GA connector, which defines how a pair of producer and consumer roles coordinates in the environment. The GA glue reconciles consumer and producer roles that may differ with respect to time and space coupling as well as scoping. Hence, GA connectors support interactions among highly heterogeneous services of the Future Internet, and especially across domains.

eXtensible Service Bus: We apply the above connector abstractions to introduce an enhanced bus paradigm, the *eXtensible Service Bus (XSB)*. XSB features richer interaction semantics than common ESB implementations to deal effectively with the increased Future Internet heterogeneity. Moreover, from its very conception, XSB incorporates special consideration for the cross-integration of heterogeneous interaction paradigms. When mapping between such paradigms, special attention is paid to the preservation of interaction semantics. XSB has the following features:

- XSB is an abstract bus that prescribes only the high-level semantics of the common bus protocol. The XSB common bus protocol features GA semantics.
- Heterogeneous systems can be plugged into the XSB by employing binding components that adapt between the native middleware of the deployed system and the common bus protocol. This adaptation is based on the systematic abstractions and mappings discussed above
- XSB, being an abstract bus, can have different implementations. This means that it needs to be complemented with a substrate which at least supports: (1) deployment (i.e., plugging) of various systems on the bus, and (2) a common bus protocol implementing GA semantics. With respect to the latter, we envision that a GA protocol realization may either be designed and built from scratch (still supposing at least an IP-based transport substrate) or be implemented by conveying GA semantics on top of an existing higher-level protocol used as transport carrier. The latter solution can be attractive, as it facilitates GA protocol realizations in different contexts and domains.

We have carried out two realizations of XSB for the CHOReOS project [30], the first on PEtALS ESB and the second on EasyESB. The genericity and modularity of our solution allowed for easily porting from the first implementation to the second one. We support interoperable peer-to-peer interaction among the CS, PS, and TS paradigms and provide templates for systematic and highly facilitated building of binding components for middleware platforms that follow any one of the three paradigms.

6.4. Service Oriented Middleware facing the Challenges of the Internet of Things

Participants: Benjamin Billet, Nikolaos Georgantas, Sara Hachem, Valérie Issarny, Yesid Jarma Alvis, Cong Kinh Nguyen, George Mathioudakis.

In our vision, The Future Internet can be defined as the union and cooperation of the Internet of Content, Internet of Services, Internet of Things, and 3D interactive Internet, supported by an expanding network infrastructure foundation [6]. In ARLES, we chose to pay special attention to the Internet of Things (IoT). IoT is characterized by the integration of large numbers of real-world objects (or “things") onto the Internet, with the aim of turning high-level interactions with the physical world into a matter as simple as is interacting with the virtual world today. As such, two devices that will play a key role in the IoT are *sensors* and *actuators*. In fact, such devices are already seeing widespread adoption in the highly localized systems within our cars, mobile phones, laptops, home appliances, etc. In their current incarnation, however, sensors and actuators are used for little more than low-level inferences and basic services. This is partly due to their highly specialized domains (signal processing, estimation theory, robotics, etc.), which demand application programmers to also be domain experts, and partly due to a glaring lack of interconnectivity between all the different devices. Our work within this domain has been focused on two related directions:

- **Architecture of a Service Oriented Middleware for the Mobile Internet of Things**: Adopting the service-oriented architecture (SOA) approach towards middleware (see § 3.3), is an adequate solution towards addressing the heterogeneity and the unknown network topology issues in the IoT. SOA is commonly used in IoT solutions to abstract *things* or their measurements as services. The service-oriented paradigm decouples the functionalities of things from their hardware information or other technical details, and supports three core functionalities: *Discovery* and *Composition* of, and *Access* to services. Typically, in traditional uses of SOA, even if millions of services are registered, there is no need to select and access them all simultaneously. However, in the IoT, discovery, composition and access are undoubtedly more complicated. In fact, it is unlikely for a single or even a few services to be sufficient when providing real world measurements. In most cases, to accurately represent a real-world feature, a large number of services are selected to provide their measurements, and subsequently, all acquired values should be properly aggregated. As a consequence, discovery will return a large set of accessible services, redundant as they may be. Consumers are then expected to access the numerous providers to acquire their measurements, over which they should know the exact aggregation/fusion logic to apply. Furthermore, such logic requires precise knowledge

and understanding of the real world and its governing physics and mathematics laws. Clearly, performing discovery, composition and access tasks as presented above incurs high communication and computation costs and is thus not realistic within the large scale IoT. In light of the above issues, we have been *revisiting the SOA and its interaction patterns* to support better scalability and exempt consumers from directly interacting with providers. Specifically, we introduce a **thing-based SOA** to wrap access and computation activities in a middleware that, unlike traditional SOA middleware, is aware of the real world, its physics and its mathematics rules; this has further led to our initial work on the components of such a middleware.

- **Probabilistic Registration for Large Scale Mobile Participatory Sensing**: An increasingly important component of the Internet of Things are modern smart phones, whose constituent sensors and wireless connectivity make them ideal candidates for *mobile participatory sensing*, which aids in providing increased knowledge about the real world while relying on a large number of mobile devices. Those devices can host different types of sensors incorporated in every aspect of our lives. However, given the increasing number of capable mobile devices, any participatory sensing approach should be, first and foremost, *scalable*. To address this challenge, we present an approach to decrease the participation of (sensing) devices in a manner that does not compromise the accuracy of the real-world information while increasing the efficiency of the overall system. To reduce the number of the devices involved, we present a probabilistic registration approach [20], based on a realistic human mobility model, that allows devices to decide whether or not to register their sensing services depending on the probability of other, equivalent devices being present at the locations of their expected path. We used our techniques as the basis of the design and implementation of a registration middleware, using which mobile devices can base their registration decision. Through experiments performed on real and simulated datasets, we show that our approach scales, while not sacrificing significant amounts of sensing coverage.

Our IoT middleware is currently being used by the industrial partners in the FP7 IP CHOReOS project. Complementary to our research on this service oriented middleware for the Internet of Things, we have also been working on suitable abstractions for enabling easy application development for the IoT, discussed next.

6.5. Composing Applications in the Internet of Things

Participants: Peter Sawyer, Pankesh Patel, Animesh Pathak.

As introduced above, the Internet of Things integrates the physical world with the existing Internet, and is rapidly gaining popularity, thanks to the increased adoption of smart phones and sensing devices. Several IoT applications have been reported in recent research, and we expect to see increased adoption of IoT concepts in the fields of personal health, inventory management, and domestic energy usage monitoring, among others.

An important challenge to be addressed in the domain of IoT is to enable domain experts (health-care professionals, architects, city planners, etc.) to develop applications in their fields rapidly, with minimal support from skilled computer science professionals. Similar challenges have already been addressed in the closely related fields of Wireless Sensor and Actuator Networks (WSANs) and Pervasive/Ubiquitous computing. While the main challenge in the former is the *large scale* of the systems (hundreds to thousands of largely similar nodes, sensing and acting on the environment), the primary concern in the latter has been the *heterogeneity* of nodes and the major role that the user's own interaction with these nodes plays in these systems (cf. the classic "smart home" scenario where the user interacts with a smart display which works together with his refrigerator and toaster). The upcoming field of IoT includes both WSANs as well as smart appliances, in addition to the elements of the "traditional" Internet such as Web and database servers, exposing their functionalities as Web services, etc. Consequently, an ideal application development abstraction of the IoT will allow (domain expert) developers to intuitively specify the rich interactions between the extremely large number of disparate devices in the future Internet of Things [19].

The larger goal of our research is to propose a suitable application development framework which addresses the challenges introduced above. To that end, our work this year covered the following related areas:

- **Multi-stage Model-driven approach for IoT Application Development**: We have proposed a multi-stage model-driven approach for IoT application development based on a precise definition of the role to be played by each stakeholder involved in the process – domain expert, application designer, application developer, device developer, and network manager. The metamodels/abstractions available to each stakeholder are further customized using the inputs provided in the earlier stages by other stakeholders. We have also implemented code-generation and task-mapping techniques to support our approach. Our initial evaluation based on two realistic scenarios shows that the use of our techniques/framework succeeds in improving productivity in the IoT application development process.
- **Revisiting Requirements Engineering (RE) Practices for IoT**: Requirements engineering (RE) has evolved to discover, model, specify and manage the required and desired properties of software systems. Conventional RE makes an assumption that the knowledge from which the requirements will be formulated exists a-priori, even though the knowledge may be fragmentary, distributed and tacit. Thus, although their discovery may take significant effort, the requirements are discoverable using the appropriate RE practices.

  However, the last decade or so has seen the emergence of new types of systems where this assumption does not hold, including the IoT. Conventional RE is ill-equipped to discover, model, specify and manage these systems' requirements because incomplete knowledge of the context under which they must operate is available at design time. While some progress has been made, by (e.g.) maintaining requirements models that support reasoning over context at runtime, the IoT has now emerged to compound the challenge for RE. Drawing on experiences from ubiquitous computing and WSAN domains, in [22] we provided initial insights into how the field of RE needs to evolve in order to address the challenges brought forth by IoT.

We have incorporated our continued research 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.

6.6. Requirements-aware Systems for Self-adaptation under Uncertainty

Participants: Romina Torres, Nelly Bencomo, Valérie Issarny, Peter Sawyer.

The development of software-intensive systems is driven by their requirements. Traditional requirements engineering (RE) methods focus on resolving ambiguities in requirements and advocate specifying requirements in sufficient detail so that the implementation can be checked against them for conformance. In an ideal world, this way of thinking can be very effective. Requirements can be specified clearly, updated as necessary, and evolutions of the software design can be made with the requirements in mind.

Increasingly, however, it is not sufficient to fix requirements statically because they will change at runtime as the operating environment changes. Furthermore, as software systems become more pervasive, there is growing uncertainty about the environment and so requirements changes cannot be predicted at design-time. It is considerations such as these that have led to the development of self-adaptive systems (SASs), which have the ability to dynamically and autonomously reconfigure their behavior to respond to changing external conditions.

The key argument of our research is that current software engineering (SE) methods do not support well the kind of dynamic appraisal of requirements needed by a SAS. definition and structure of requirements is lost as requirements are refined into an implementation. Even in cases where requirements monitoring is explicitly included, high-level system requirements must be manually refined into low-level runtime artefacts during the design process so that they can be monitored. There is a lack of approaches supporting for runtime representation, evolution and assessment of requirements. Currently, the approaches mainly assume that it is possible to predefine and envisage the requirements for the total set of target behaviours. Such estimations

and beliefs may not be appropriate, if the system is to recover during execution from unforeseen situations, or adapt dynamically to new environmental conditions or to satisfy new requirements that were not foreseen during development. A self-adaptive system is able, at run time, to satisfy new requirements and behaviors. Our research focuses on approaches to support the runtime representation of requirements that will underpin the way a system can reason and assess them during execution.

Our research has been carried out within the research project Marie Curie Fellowship called Requirements-aware Systems (nickname: Requirements@run.time). The research is based on a new paradigm for SE, called requirements-awareness (also known as requirements reflection), in which requirements are reified as runtime entities. Requirements-awareness allows systems to dynamically reason about themselves at the level of the requirements - in much the same way that architectural reflection currently allows runtime reasoning at the level of software. We believe that requirements-awareness (i.e. requirements reflection) will support the development and management of SASs because it will raise the level of discourse at which a software system is able to reflect upon itself.

In the above context, we have been working on the design and implementation of systems with the ability to dynamically observe and reason about their requirements. The results will contribute towards the development of conceptual foundations, engineering techniques, and computing infrastructure for the access and manipulations of runtime abstractions of requirements. Currently, a prototype for the use of runtime goals has been developed. The RELAX language has been proposed to make requirements more tolerant to environmental uncertainty. Design assumptions, called Claims), are applied as markers of uncertainty that document how design assumptions affect goals. Monitoring Claims at runtime has been used to drive self-adaptation. By monitoring Claims during the execution of the systems, their veracity can be tested. If a Claim is falsified, the effect can be propagated to the system's goal model and an alternative (more suitable) means of goal realization will be selected, resulting in dynamic adaptation of the system to a configuration that better satisfies the goals under the prevailing environmental context.

GANG Project-Team

4. New Results

4.1. Understanding graph representations

4.1.1. Notions of Connectivity in Overlay Networks

Participants: Yuval Emek, Pierre Fraigniaud, Amos Korman, Shay Kutten, David Peleg.

How well connected is the network? This is one of the most fundamental questions one would ask when facing the challenge of designing a communication network. Three major notions of connectivity have been considered in the literature, but in the context of traditional (single-layer) networks, they turn out to be equivalent. The paper [17], introduces a model for studying the three notions of connectivity in multi-layer networks. Using this model, it is easy to demonstrate that in multi-layer networks the three notions may differ dramatically. Unfortunately, in contrast to the single-layer case, where the values of the three connectivity notions can be computed efficiently, it has been recently shown in the context of WDM networks (results that can be easily translated to our model) that the values of two of these notions of connectivity are hard to compute or even approximate in multi-layer networks. The current paper shed some positive light into the multi-layer connectivity topic: we show that the value of the third connectivity notion can be computed in polynomial time and develop an approximation for the construction of well connected overlay networks.

4.1.2. Connected graph searching

Participants: Lali Barrière, Paola Flocchini, Fedor V. Fomin, Pierre Fraigniaud, Nicolas Nisse, Nicola Santoro, Dimitrios M. Thilikos.

In the graph searching game the opponents are a set of searchers and a fugitive in a graph. The searchers try to capture the fugitive by applying some sequence of moves that include placement, removal, or sliding of a searcher along an edge. The fugitive tries to avoid capture by moving along unguarded paths. The search number of a graph is the minimum number of searchers required to guarantee the capture of the fugitive. In [2], we initiate the study of this game under the natural restriction of connectivity where we demand that in each step of the search the locations of the graph that are clean (i.e. non-accessible to the fugitive) remain connected. We give evidence that many of the standard mathematical tools used so far in classic graph searching fail under the connectivity requirement. We also settle the question on "the price of connectivity", that is, how many searchers more are required for searching a graph when the connectivity demand is imposed. We make estimations of the price of connectivity on general graphs and we provide tight bounds for the case of trees. In particular, for an n-vertex graph the ratio between the connected searching number and the non-connected one is while for trees this ratio is always at most 2. We also conjecture that this constant-ratio upper bound for trees holds also for all graphs. Our combinatorial results imply a complete characterization of connected graph searching on trees. It is based on a forbidden-graph characterization of the connected search number. We prove that the connected search game is monotone for trees, i.e. restricting search strategies to only those where the clean territories increase monotonically does not require more searchers. A consequence of our results is that the connected search number can be computed in polynomial time on trees, moreover, we show how to make this algorithm distributed. Finally, we reveal connections of this parameter to other invariants on trees such as the Horton–Strahler number.

4.1.3. Computing with Large Populations Using Interactions

Participants: Olivier Bournez, Pierre Fraigniaud, Xavier Koegler.

We define in [12], a general model capturing the behavior of a population of anonymous agents that interact in pairs. This model captures some of the main features of opportunistic networks, in which nodes (such as the ones of a mobile ad hoc networks) meet sporadically. For its reminiscence to Population Protocol, we call our model Large-Population Protocol, or LPP. We are interested in the design of LPPs enforcing, for every $\nu \in [0,1]$, a proportion ν of the agents to be in a specific subset of marked states, when the size of the population grows to infinity; In which case, we say that the protocol computes ν. We prove that, for every $\nu \in [0,1]$, ν is computable by a LPP if and only if ν is algebraic. Our positive result is constructive. That is, we show how to construct, for every algebraic number $\nu \in [0,1]$, a protocol which computes ν.

4.1.4. Collaborative Search on the Plane without Communication

Participants: Ofer Feinerman, Zvi Lotker, Amos Korman, Jean-Sébastien Sereni.

In [19], we use distributed computing tools to provide a new perspective on the behavior of cooperative biological ensembles. We introduce the Ants Nearby Treasure Search (ANTS) problem, a generalization of the classical cow-path problem which is relevant for collective foraging in animal groups. In the ANTS problem, k identical (probabilistic) agents, initially placed at some central location, collectively search for a treasure in the two-dimensional plane. The treasure is placed at a target location by an adversary and the goal is to find it as fast as possible as a function of both k and D, where D is the distance between the central location and the target. This is biologically motivated by cooperative, central place foraging, such as performed by ants around their nest. In this type of search there is a strong preference to locate nearby food sources before those that are further away. We focus on trying to find what can be achieved if communication is limited or altogether absent. Indeed, to avoid overlaps agents must be highly dispersed making communication difficult. Furthermore, if the agents do not commence the search in synchrony, then even initial communication is problematic. This holds, in particular, with respect to the question of whether the agents can communicate and conclude their total number, k. It turns out that the knowledge of k by the individual agents is crucial for performance. Indeed, it is a straightforward observation that the time required for finding the treasure is $\Omega(D + D^2/k)$, and we show in this paper that this bound can be matched if the agents have knowledge of k up to some constant approximation. We present a tight bound for the competitive penalty that must be paid, in the running time, if the agents have no information about k. Specifically, this bound is slightly more than logarithmic in the number of agents. In addition, we give a lower bound for the setting in which the agents are given some estimation of k. Informally, our results imply that the agents can potentially perform well without any knowledge of their total number k, however, to further improve, they must use some information regarding k. Finally, we propose a uniform algorithm that is both efficient and extremely simple, suggesting its relevance for actual biological scenarios.

4.1.5. Memory Lower Bounds for Randomized Collaborative Search and Implications for Biology

Participants: Ofer Feinerman, Amos Korman.

Initial knowledge regarding group size can be crucial for collective performance. We study in [18], this relation in the context of the Ants Nearby Treasure Search (ANTS) problem, which models natural cooperative foraging behavior such as that performed by ants around their nest. In this problem, k (probabilistic) agents, initially placed at some central location, collectively search for a treasure on the two-dimensional grid. The treasure is placed at a target location by an adversary and the goal is to find it as fast as possible as a function of both k and D, where D is the (unknown) distance between the central location and the target. It is easy to see that $T = \Omega(D + D^2/k)$ time units are necessary for finding the treasure. Recently, it has been established that $O(T)$ time is sufficient if the agents know their total number k (or a constant approximation of it), and enough memory bits are available at their disposal. In this paper, we establish lower bounds on the agent memory size required for achieving certain running time performances. To the best our knowledge, these bounds are the first non-trivial lower bounds for the memory size of probabilistic searchers. For example, for every given positive constant ϵ, terminating the search by time $O(\log^{1-\epsilon} k \cdot T)$ requires agents to use $\Omega(\log \log k)$ memory bits.

From a high level perspective, we illustrate how methods from distributed computing can be useful in generating lower bounds for cooperative biological ensembles. Indeed, if experiments that comply with our setting reveal that the ants' search is time efficient, then our theoretical lower bounds can provide some insight on the memory they use for this task.

4.1.6. What Can be Computed without Communications?

Participants: Heger Arfaoui, Pierre Fraigniaud.

When playing the boolean game (δ, f), two players, upon reception of respective inputs x and y, must respectively output a and b satisfying $\delta(a, b) = f(x, y)$, in absence of any communication. It is known that, for $\delta(a, b) = a \oplus b$, the ability for the players to use entangled quantum bits (qbits) helps. In [10], we show that, for δ different from the exclusive-or operator, quantum correlations do not help. This result is an invitation to revisit the theory of dis- tributed checking, a.k.a. distributed verification, currently sticked to the usage of decision functions δ based on the and-operator, hence potentially preventing us from using the potential benefit of quantum effects.

4.1.7. Decidability Classes for Mobile Agents Computing modularity

Participants: Andrzej Pelc, Pierre Fraigniaud.

We establish in [21], a classification of decision problems that are to be solved by mobile agents operating in unlabeled graphs, using a deterministic protocol. The classification is with respect to the ability of a team of agents to solve the problem, possibly with the aid of additional information. In particular, our focus is on studying differences between the decidability of a decision problem by agents and its verifiability when a certificate for a positive answer is provided to the agents. Our main result shows that there exists a natural complete problem for mobile agent verification. We also show that, for a single agent, three natural oracles yield a strictly increasing chain of relative decidability classes.

4.1.8. Randomized Distributed Decision

Participants: Pierre Fraigniaud, Amos Korman, Merav Parter, David Peleg.

The paper [20] tackles 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.

We focus on the notion of a $(p, q)-$decider for a language L , which is a distributed randomized algorithm that accepts instances in L with probability at least p and rejects instances outside of L with probability at least q. It is known that every hereditary language that can be decided in t rounds by a $(p, q)-$decider, where $p^2 + q > 1$, can be decided deterministically in $O(t)$ rounds. One of our results gives evidence supporting the conjecture that the above statement holds for all distributed languages and not only for hereditary ones, by proving the conjecture for the restricted case of path topologies. For the range below the aforementioned threshold, namely, $p^2 + q \leq 1$, we study the class $B_k(t)$ (for $k \in \mathbb{N}^* \cup \{\infty\}$) of all languages decidable in at most t rounds by a $(p, q)-$decider, where $p^{1+\frac{1}{k}} + q > 1$. Since every language is decidable (in zero rounds) by a $(p, q)-$decider satisfying $p + q = 1$, the hierarchy B_k provides a spectrum of complexity classes between determinism ($k = 1$, under the above conjecture) and complete randomization ($k = \infty$). We prove that all these classes are separated, in a strong sense: for every integer $k \geq 1$, there exists a language L satisfying $L \in B_{k+1}(0)$ but $L \notin B_k(t)$ for any $t = o(n)$. In addition, we show that $B_\infty(t)$ does not contain all languages, for any $t = o(n)$. In other words, we obtain the hierarchy $B_(t) \subset B_2(t) \subset \cdots \subset B_\infty(t) \subset$ All. Finally, we show that if the inputs can be restricted in certain ways, then the ability to boost the success probability becomes almost null, and in particular, derandomization is not possible even beyond the threshold $p^2 + q = 1$.

4.1.9. The Worst Case Behavior of Randomized Gossip

Participants: Hervé Baumann, Pierre Fraigniaud, Hovhannes A. Harutyunyan, Rémi de Verclos.

In [11] we consider the quasi-random rumor spreading model introduced by Doerr, Friedrich, and Sauerwald in [SODA 2008], hereafter referred to as the list-based model. Each node is provided with a cyclic list of all its neighbors, chooses a random position in its list, and from then on calls its neighbors in the order of the list. This model is known to perform asymptotically at least as well as the random phone-call model, for many network classes. Motivated by potential applications of the list-based model to live streaming, we are interested in its worst case behavior.

Our first main result is the design of an $O(m + nlogn)-$time algorithm that, given any n-node m-edge network G, and any source-target pair $s, t \in V(G)$, computes the maximum number of rounds it may take for a rumor to be broadcast from s to t in G, in the list-based model. This algorithm yields an $O(n(m + nlogn))-$time algorithm that, given any network G, computes the maximum number of rounds it may take for a rumor to be broadcast from any source to any target, in the list-based model. Hence, the list-based model is computationally easy to tackle in its basic version.

The situation is radically different when one is considering variants of the model in which nodes are aware of the status of their neighbors, i.e., are aware of whether or not they have already received the rumor, at any point in time. Indeed, our second main result states that, unless P=NP , the worst case behavior of the list-based model with the additional feature that every node is perpetually aware of which of its neighbors have already received the rumor cannot be approximated in polynomial time within a $\left(\frac{1}{n}\right)^{\frac{1}{2}-\epsilon}$ multiplicative factor, for any $\epsilon > 0$. As a byproduct of this latter result, we can show that, unless P=NP , there are no PTAS enabling to approximate the worst case behavior of the list-based model, whenever every node perpetually keeps track of the subset of its neighbors which have sent the rumor to it so far.

4.1.10. Asymptotic modularity

Participants: Fabien de Montgolfier, Mauricio Soto, Laurent Viennot.

Modularity (Newman-Girvan) has been introduced as a quality measure for graph partitioning. It has received considerable attention in several disciplines, especially complex systems. In order to better understand this measure from a graph theoretical point of view, we study the modularity of a variety of graph classes. In [23], we first consider simple graph classes such as tori and hypercubes. We show that these regular graph families have asymptotic modularity 1 (that is the maximum possible). We extend this result to trees with bounded degree, allowing us to give a lower bound of 2 over average degree for graph classes with low maximum degree (included power law graphs for a sufficiently large exponent).

4.1.11. Modeling social networks

Participants: Nidhi Hegde, Laurent Massoulié, Laurent Viennot.

Social networks offer users new means of accessing information, essentially relying on “social filtering”, i.e. propagation and filtering of information by social contacts. The sheer amount of data flowing in these networks, combined with the limited budget of attention of each user, makes it difficult to ensure that social filtering brings relevant content to the interested users. Our motivation in [26] is to measure to what extent self-organization of the social network results in efficient social filtering. To this end we introduce flow games, a simple abstraction that models network formation under selfish user dynamics, featuring user-specific interests and budget of attention. In the context of homogeneous user interests, we show that selfish dynamics converge to a stable network structure (namely a pure Nash equilibrium) with close-to-optimal information dissemination. We show in contrast, for the more realistic case of heterogeneous interests, that convergence, if it occurs, may lead to information dissemination that can be arbitrarily inefficient, as captured by an unbounded “price of anarchy”. Nevertheless the situation differs when users’ interests exhibit a particular structure, captured by a metric space with low doubling dimension. In that case, natural autonomous dynamics converge to a stable configuration. Moreover, users obtain all the information of interest to them in the corresponding dissemination, provided their budget of attention is logarithmic in the size of their interest set.

4.1.12. Additive Spanners and Distance and Routing Labeling Schemes for Hyperbolic Graphs

Participants: Victor Chepoi, Feodor Dragan, Bertrand Estellon, Michel Habib, Yann Vaxès, Yang Xiang.

δ-Hyperbolic metric spaces have been defined by M. Gromov in 1987 via a simple 4-point condition: for any four points u, v, w, x, the two larger of the distance sums $d(u,v)+d(w,x)$, $d(u,w)+d(v,x)$, $d(u,x)+d(v,w)$ differ by at most 2δ. They play an important role in geometric group theory, geometry of negatively curved spaces, and have recently become of interest in several domains of computer science, including algorithms and networking. In [5], we study unweighted δ-hyperbolic graphs. Using the Layering Partition technique, we show that every n-vertex δ-hyperbolic graph with $\delta \geq 1/2$ has an additive $O(\delta \log n)$-spanner with at most $O(\delta n)$ edges and provide a simpler, in our opinion, and faster construction of distance approximating trees of δ-hyperbolic graphs with an additive error $O(\delta \log n)$. The construction of our tree takes only linear time in the size of the input graph. As a consequence, we show that the family of n-vertex δ-hyperbolic graphs with $\delta \geq$1/2 admits a routing labeling scheme with $O(\delta \log^2 n)$ bit labels, $O(\delta \log n)$ additive stretch and $O(\log_2(4\delta))$ time routing protocol, and a distance labeling scheme with $O(\log^2 n)$ bit labels, $O(\delta \log n)$ additive error and constant time distance decoder.

4.1.13. Constructing a Minimum phylogenetic Network from a Dense triplet Set

Participants: Michel Habib, Thu-Hien To.

For a given set $\mathcal{L}$ of species and a set $\mathcal{T}$ of triplets on $\mathcal{L}$, we seek to construct a phylogenetic network which is consistent with $\mathcal{T}$ i.e. which represents all triplets of $\mathcal{T}$. The level of a network is defined as the maximum number of hybrid vertices in its biconnected components. When $\mathcal{T}$ is dense, there exist polynomial time algorithms to construct level-0, 1 and 2 networks (Aho et al., 1981; Jansson, Nguyen and Sung, 2006; Jansson and Sung, 2006; Iersel et al., 2009). For higher levels, partial answers were obtained in the paper by Iersel and Kelk (2008), with a polynomial time algorithm for simple networks. In [9] this paper, we detail the first complete answer for the general case, solving a problem proposed in Jansson and Sung (2006) and Iersel et al. (2009). For any k fixed, it is possible to construct a level-k network having the minimum number of hybrid vertices and consistent with $\mathcal{T}$, if there is any, in time $O\left(|(T)|^{k+1} n^{\lfloor \frac{4k}{3} \rfloor}\right)$.

4.1.14. Algorithms for Some $H-$Join Decompositions

Participants: Michel Habib, Antoine Mamcarz, Fabien de Montgolfier.

A homogeneous pair (also known as a 2-module) of a graph is a pair $\{M_1, M_2\}$ of disjoint vertex subsets such that for every vertex $x \notin (M_1 \cup M_2)$ and $i \in \{1,2\}$, x is either adjacent to all vertices in M_i or to none of them. First used in the context of perfect graphs [Chvátal and Sbihi 1987], it is a generalization of splits (a.k.a 1-joins) and of modules. The algorithmics to compute them appears quite involved. In [22], we describe an $O(mn^2)-$time algorithm computing (if any) a homogeneous pair, which not only improves a previous bound of $O(mn^3)$ [Everett, Klein and Reed 1997], but also uses a nice structural property of homogenous pairs. Our result can be extended to compute the whole homogeneous pair decomposition tree, within the same complexity. Using similar ideas, we present an $O(nm^2)-$time algorithm to compute a $N-$join decomposition of a graph, improving a previous $O(n^6)$ algorithm [Feder et al. 2005]. These two decompositions are special case of $H-$joins [Bui-Xuan, Telle and Vatshelle 2010] to which our techniques apply.

4.1.15. Detecting 2-joins faster

Participants: Pierre Charbit, Michel Habib, Nicolas Trotignon, Kristina Vušković.

2-joins are edge cutsets that naturally appear in the decomposition of several classes of graphs closed under taking induced subgraphs, such as balanced bipartite graphs, even-hole-free graphs, perfect graphs and claw-free graphs. Their detection is needed in several algorithms, and is the slowest step for some of them. The classical method to detect a 2-join takes $O(n^3 m)$ time where n is the number of vertices of the input graph and m the number of its edges. To detect *non-path* 2-joins (special kinds of 2-joins that are needed in all of the known algorithms that use 2-joins), the fastest known method takes time $O(n^4 m)$. Here, we give an $O(n^2 m)$-time algorithm for both of these problems. A consequence is a speed up of several known algorithms.

4.2. Large Scale Networks Performance and Modeling

4.2.1. Spatial Interactions of Peers and Performance of File Sharing Systems

Participants: François Baccelli, Fabien Mathieu, Ilkka Norros.

We propose in [24] a new model for peer-to-peer networking which takes the network bottlenecks into account beyond the access. This model allows one to cope with key features of P2P networking like degree or locality constraints or the fact that distant peers often have a smaller rate than nearby peers. We show that the spatial point process describing peers in their steady state then exhibits an interesting repulsion phenomenon. We analyze two asymptotic regimes of the peer-to-peer network: the fluid regime and the hard–core regime. We get closed form expressions for the mean (and in some cases the law) of the peer latency and the download rate obtained by a peer as well as for the spatial density of peers in the steady state of each regime, as well as an accurate approximation that holds for all regimes. The analytical results are based on a mix of mathematical analysis and dimensional analysis and have important design implications. The first of them is the existence of a setting where the equilibrium mean latency is a decreasing function of the load, a phenomenon that we call super-scalability.

4.2.2. User Behavior Modeling: Four Months in DailyMotion

Participants: Yannick Carlinet, The Dang Huynh, Bruno Kauffmann, Fabien Mathieu, Ludovic Noirie, Sébastien Tixeuil.

The growth of User-Generated Content (UGC) traffic makes the understanding of its nature a priority for network operators, content providers and equipment suppliers. In [13], we study a four-month dataset that logs all video requests to DailyMotion made by a fixed subset of users. We were able to infer user sessions from raw data, to propose a Markovian model of these sessions, and to study video popularity and its evolution over time. The presented results are a first step for synthesizing an artificial (but realistic) traffic that could be used in simulations or experimental testbeds.

4.2.3. Multi-Carrier Networks: on the Manipulability of Voting Systems

Participants: François Durand, Fabien Mathieu, Ludovic Noirie.

Today, Internet involves many actors who are making revenues on it (operators, companies, service providers,...). It is therefore important to be able to make fair decisions in this large-scale and highly competitive economical ecosystem. One of the main issues is to prevent actors from manipulating the natural outcome of the decision process. For that purpose, game theory is a natural framework. In that context, voting systems represent an interesting alternative that, to our knowledge, has not yet been considered. They allow competing entities to decide among different options. Strong theoretical results showed that all voting systems are susceptible to be manipulated by one single voter, except for some "degenerated" and non-acceptable cases. However, very little is known about how much a voting system is manipulable in practical scenarios. In [25], we investigate empirically the use of voting systems for choosing end-to-end paths in multi-carrier networks, analyzing their manipulability and their economical efficiency. We show that one particular system, called Single Transferable Vote (STV), is largely more resistant to manipulability than the natural system which tries to get the economical optimum. Moreover, STV manages to select paths close to the economical optimum, whether the participants try to cheat or not.

4.3. Fault Tolerance in Distributed Networks

4.3.1. Wait-Freedom with Advice

Participants: Carole Delporte-Gallet, Hugues Fauconnier, Eli Gafni, Petr Kuznetsov.

In [14], 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 output in a finite number of steps and *synchronization* processes that are subject to failures and can query a failure detector. We assume that, 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.

4.3.2. Partial synchrony based on set timeliness

Participants: Markos Aguilera, Carole Delporte-Gallet, Hugues Fauconnier, Sam Toueg.

We introduce in [1], a new model of partial synchrony for read-write shared memory systems. This model is based on the simple notion of set timeliness—a natural generalization of the seminal concept of timeliness in the partially synchrony model of Dwork et al. (J. ACM 35(2):288–323, 1988). Despite its simplicity, the concept of set timeliness is powerful enough to define a family of partially synchronous systems that closely match individual instances of the $t-$resilient $k-$set agreement problem among n processes, henceforth denoted $(t, k, n)-$agreement. In particular, we use it to give a partially synchronous system that is synchronous enough for solving $(t, k, n)-$ agreement, but not enough for solving two incrementally stronger problems, namely, $(t + 1, k, n)-$agreement, which has a slightly stronger resiliency requirement, and $(t, k-1, n)-$agreement, which has a slightly stronger agreement requirement. This is the first partially synchronous system that separates these sub-consensus problems. The above results show that set timeliness can be used to study and compare the partial synchrony requirements of problems that are strictly weaker than consensus.

4.3.3. Byzantine Agreement with Homonyms in Synchronous Systems

Participants: Carole Delporte-Gallet, Hugues Fauconnier, Hung Tran-The.

In [15], [6], we consider the Byzantine agreement problem (BA) in synchronous systems with homonyms. In this model different processes may have the same authenticated identifier. In such a system of n processes sharing a set of l identifiers, we define a distribution of the identifiers as an integer partition of n into l parts $n_1, ..., n_l$ giving for each identifier i the number of processes having this identifier.

Assuming that the processes know the distribution of identifiers we give a necessary and sufficient condition on the integer partition of n to solve the Byzantine agreement with at most t Byzantine processes. Moreover we prove that there exists a distribution of l identifiers enabling to solve Byzantine agreement with at most t Byzantine processes if and only if $n > 3t$, $l > t$ and where $r = n \mod l$.

This bound is to be compared with the $l > 3t$ bound proved in Delporte-Gallet et al. (2011) when the processes do not know the distribution of identifiers.

4.3.4. Homonyms with forgeable identifiers

Participants: Carole Delporte-Gallet, Hugues Fauconnier, Hung Tran-The.

In [16], we refine the Byzantine Agreement problem (BA) in synchronous systems with homonyms, in the particular case where some identifiers may be forgeable. More precisely, the n processes share a set of l $(1 \leq l \leq n)$ identifiers. Assuming that at most t processes may be Byzantine and at most k $(t \leq k \leq l)$ of these identifiers are forgeable in the sense that any Byzantine process can falsely use them, we prove that Byzantine Agreement problem is solvable if and only if $l > 2t + k$. Moreover we extend this result to systems with authentication by signatures in which at most k signatures are forgeable and we prove that Byzantine Agreement problem is solvable if and only if $l > t + k$.

4.4. Discrete Optimization Algorithms

4.4.1. Estimating satisfiability

Participants: Yacine Boufkhad, Thomas Hugel.

The problem of estimating the proportion of satisfiable instances of a given CSP (constraint satisfaction problem) can be tackled through weighting. It consists in putting onto each solution a non-negative real value based on its neighborhood in a way that the total weight is at least 1 for each satisfiable instance. We define in [3], a general weighting scheme for the estimation of satisfiability of general CSPs. First we give some sufficient conditions for a weighting system to be correct. Then we show that this scheme allows for an improvement on the upper bound on the existence of non-trivial cores in 3-SAT obtained by Maneva and Sinclair (2008) to 4.419. Another more common way of estimating satisfiability is ordering. This consists in putting a total order on the domain, which induces an orientation between neighboring solutions in a way that prevents circuits from appearing, and then counting only minimal elements. We compare ordering and weighting under various conditions.

4.4.2. Attractive force search algorithm for piecewise convex maximization problems

Participants: Dominique Fortin, Ider Tseveendorj.

In [8], we consider mathematical programming problems with the so-called piecewise convex objective functions. A solution method for this interesting and important class of nonconvex problems is presented. This method is based on Newton's law of universal gravitation, multicriteria optimization and Helly's theorem on convex bodies. Numerical experiments using well known classes of test problems on piecewise convex maximization, convex maximization as well as the maximum clique problem show the efficiency of the approach.

4.4.3. B-spline interpolation: Toeplitz inverse under corner perturbations

Participant: Dominique Fortin.

For Toeplitz matrices associated with degree 3 and 4 uniform B-spline interpolation, the inverse may be analytically known [7], saving the standard inverse calculations. It generalizes to any degree as a row of the Eulerian numbers triangle.

HIPERCOM Project-Team

6. New Results

6.1. Time Slot Assignment in Wireless Sensor Networks

Participants: Pascale Minet, Ridha Soua, Erwan Livolant.

6.1.1. NP-completeness of the Time Slot Assignment problem

In data gathering applications, wireless sensor networks (WSNs) collect data from sensor nodes towards a sink in a multi-hop convergecast structure. Assigning equal channel access to each node may lead to congestion and inefficient use of the bandwidth. That is why we focus on traffic-aware solutions. More precisely, we investigate the Time Slot Assignment problem, where nodes are assigned time slots to transmit their data to the sink, while minimizing the total number of slots. We considered the generalized h-hop Time Slot Assignment problem for any positive integer h, where any two nodes that are less than or equal to h-hop away are not scheduled simultaneously. We proved its NP-completeness.

6.1.2. Multichannel Slot Assignment

The throughput requirement of data gathering applications is difficult to meet with a single wireless channel. Furthermore, the considered channel may be temporarily jammed. That is why, we focus on a multichannel time slot assignment that minimizes the data gathering cycle. We first formalize the problem as a linear program and compute the optimal time needed for a raw data convergecast in various multichannel topologies (linear, multi-line, tree). These optimal times apply to nodes equipped with one or several radio interfaces. This work generalizes the results established by Incel. We then propose our algorithm called MODESA and prove its optimality in various multichannel topologies. We evaluate its performances in terms of number of slots, maximum buffer size and number of active/sleep switches per node. Furthermore, we present variants of MODESA achieving a load balancing between the channels used.

6.1.3. Multisink Multichannel Slot Assignment

We generalize this work, taking into account the existence of several sinks. We focus on the data gathering problem with differentiated traffic, each addressed to a specific sink in multichannel WSNs. In order to find a collision-free optimized multichannel time slot assignment that minimizes the data gathering cycle, we propose a centralized traffic-aware algorithm called MUSIKA. We formulate the problem as a linear program and compute the optimal time needed for a raw data convergecast in various multichannel topologies (linear, multi-line, tree). More generally, we run simulations on various network topologies to evaluate the performance of MUSIKA in terms of cycle length, maximum buffer size and slot reuse ratio for different use cases: redundant functional processing chains, different application functionalities per sink.

6.2. Multi-Sink Wireless Sensor deployment and energy analysis

Participants: Paul Mühlethaler, Nadjib Achir.

We propose a general framework for multi-sink Wireless Sensors networks (WNSs). This framework is devoted to computing the optimal deployment of sinks for a given maximum number of hops between nodes and sinks. This framework allows an estimation of the energy consumption to be computed. 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. We determine the upper-bound and lower-bound of overhearing. We also propose another estimation which can 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 paper are obtained using ILP formalization.

6.3. WSN Redeployment

Participants: Pascale Minet, Saoucene Mahfoudh Ridene, Ines Khoufi.

This is a joint work with Telecom SudParis: Anis Laouiti.

6.3.1. Centralized redeployment algorithm based on Virtual Forces

In many applications (e.g military, environment monitoring), wireless sensors are randomly deployed in a given area. Unfortunately, this deployment is not efficient enough to ensure full area coverage and total network connectivity. Hence, all the considered area must be covered by sensors ensuring that any event is detected in the sensing range of at least one sensor. In addition, the sensor network must be connected in terms of radio communication in order to forward the detected event to the sink(s). Thus, a redeployment algorithm has to be applied in order to achieve these two goals. In this context, we have proposed redeployment algorithms based on virtual forces. First, we have designed and simulated a centralized algorithm called CVFA. This algorithm is executed by a specific node which has global information of node positions.

6.3.2. Distributed redeployment algorithm based on Virtual Forces

Then, we proposed DVFA, Distributed Virtual forces Algorithm. Each node in the network executes DVFA and computes its new position based on information collected from its neighbors.
Performance evaluation shows that both CVFA and DVFA give very good coverage rate (between 98% and 100%) and ensure the connectivity between sensors.

6.3.3. Distributed redeployment algorithm based on Virtual Forces in the presence of obstacles

Moreover, in a real environment, obstacles such as trees, walls and buildings may exist and they may impact the deployment of wireless sensors. Obstacles can prohibit the network connectivity between nodes and create some uncovered holes or some accumulation of sensors in the same region. Consequently, an efficient wireless sensors deployment algorithm is required to ensure both coverage and network connectivity in the presence of obstacles. We have focused on this problem and enhanced our Distributed Virtual Force Algorithm (DVFA) to cope with obstacles. Simulation results show that DVFA gives very good performances even in the presence of obstacles.

6.4. Mesh Network Planning: Deployment and Canal Allocation

Participant: Nadjib Achir.

This is a joint work with University Paris XIII: A. Farsi, K. Boussetta.
We deal with the Wireless LAN planning problem. We study this problem and we propose to couple its two major issues: AP placement and channel assignment to treat them jointly. Here, we propose a novel fast and scalable three-phase heuristic algorithm (TPHA). Our proposal is able to resolve the defined multiobjective problem to provide (1) the efficient number of Access Points (APs) to be deployed, while (2) ensuring the coverage of all Test Points (TPs) and (3) maximizing their nominal data rate. To achieve the first objective, we propose an heuristic called MCL-ILP combining the quick decision making based on the Markovian CLustering algorithm and the exact solution provided by the Integer Linear Programming. Hence, a TPs-based Least Interfering Channel Search algorithm (TLICS) has been proposed for channel assignment to improve the throughput at TP locations. However, the Virtual Forces-based WLAN Planning Algorithm namely VFPA considers the results delivered by the two previous algorithms as an initial solution and tries to enhance it by adjusting the APs' positions and re-assigning their operating frequencies. Computational results exhibit that our proposal is highly beneficial to designing WLANs.

6.5. Routing in MANETs using slotted Aloha. End-to-end delays

Participants: Paul Mühlethaler, Iskander Banaouas.

This is a joint work with TREC: B. Blaszczyzyn.

Planar Poisson models with the Aloha medium access scheme have already proved to be very useful in studies of mobile ad-hoc networks (MANETs). However, it seems difficult to quantitatively study the performances of end-to-end routing in these models. In order to tackle this problem, in this paper we study a *linear stationary route embedded in an independent planar field of interfering nodes*. We consider this route as an idealization of a "typical" route in a MANET obtained by some routing mechanism. Such a decoupling allows us to obtain many numerically tractable expressions for local and mean end-to-end delays and the speed of packet progression, assuming slotted Aloha MAC and the Signal-to-Interference-and-Noise Ratio (SINR) capture condition, with the usual power-law path loss model and Rayleigh fading. These expressions show how the network performance depends on the tuning of Aloha and routing parameters and on the external noise level. In particular we show a need for a well-tuned lattice structure of fixed relaying nodes, which helps to relay packets on long random routes in the presence of a non-negligible noise. We also consider a *Poisson-line MANET model*, in which *all* nodes are located on roads forming a Poisson-line process. In this case our linear route is rigorously (in the sense of Palm theory) the typical route in this Poisson-line MANET.

6.6. Cognitive networks using a darwinian approach

Participant: Paul Mühlethaler.

This is a joint work with Alcatel Bell Labs: Philippe Jacquet.
We present a new approach for cognitive radio. In the usual approach the secondary network is in charge of monitoring the channel to determine whether or not the primary network is active in the area. If it is not, the secondary network is allowed to use the spectrum. In the new access scheme we propose, the primary network encompasses the techniques which allow it to capture the bandwidth even if the secondary network is transmitting in the area. The access scheme of the primary network preempts the secondary network activity. We present an access scheme which preempts the IEEE 802.11 decentralized scheme. This protocol is a generalized Carrier Sense Multiple Access scheme using active signaling. Instead of only sensing the carrier, this algorithm also transmits bursts of signal which may be sensed by the other nodes. If so, they give up the selection process. We show that this scheme preempts the IEEE 802.11 decentralized access scheme if the bursts transmitted by the node in the primary network are made up of special sequences which alternate between bursts of signal and periods of sensing. These sequences called (d, k) sequences encompass a minimum number d and a maximum number of k successive zeros during which the node senses the channel to find other possible concurrent transmissions. In practice we use $d = 0$ and k depends on the duration of the IEEE 802.11 interframe space and the duration of a signaling burst. We compute the number of $(0, k)$ sequences with respect to the length n of the sequence. We also show that (d, k) sequences (with $2d > k$) can be used if, by mistake, during the signaling phase one burst is not detected. We evaluate the number of such sequences.

6.7. Massive mobile dense wireless networks

Participants: Aline Carneiro Viana, Ana Cristina B. Kochem Vendramin, Kanchana Thilakarathna, Eduardo Mucceli.

routing protocols, analytical models, content distribution.

6.7.1. Scientific achievements

6.7.1.1. Social Relationship Classified

Understanding human mobility is of fundamental importance when designing new communication protocols that exploit opportunistic encounters among users. In particular, human behavior is characterized by an elevated rate of regularity, but random events are always possible in the routines of individuals as hardly predictable situations that deviate from the regular pattern and are unlikely to arise repeatedly in the future. These random events veil the ordinary patterns by introducing a significant amount of noise, thus making the process of knowledge discovery in social dataset a complex task. However, the ability to accurately identify random and social events in large datasets is essential to social analysis as well as to applications that rely

on a precise description of human routines, such as recommendation systems, forwarding strategies and ad-hoc message dissemination schemes focusing on coverage efficiency with a limited number of redundant messages. In such a context, we have proposed a strategy to analyze wireless network scenarios where mobile users interact in a rational manner, reflecting their interests and activity dynamics. Our strategy, named Random rElationship ClASsifier sTrategy (RECAST), allows to classify user relationships, separating random interactions from different kinds of social ties. The goal is achieved by observing how the real system differs from an equivalent one where entities decisions are completely random. We have evaluate the effectiveness of RECAST classification on datasets of real-world user contacts in diverse networking contexts. Our analysis unveils significant differences in the relationship dynamics of the datasets, proving that the evaluation of network protocols on a single dataset cannot lead to conclusions of general validity.

6.7.1.2. Social-aware Forwarding Protocol

Pervasiveness of computing devices, ubiquitous wireless communication, emergence of new applications, and cloud services are examples of current new emerging factors that emphasize the increasing need for adaptive networking solutions. The adaptation, most of the time, requires the design of more interdisciplinary approaches as those inspired by techniques coming from biology, social structures, games, and control systems. The approach we consider brings together solutions from different but complementary domains - i.e., networking, biology, and complex networks - aiming to deal with the problem of efficient data delivery in mobile and intermittently connected networks. For this, we have designed the Cultural Greedy Ant (CGrAnt) protocol to solve the problem of data delivery in mobile and intermittently connected networks referred as Delay Tolerant Networks (DTNs). CGrAnt is a hybrid Swarm Intelligence-based forwarding protocol designed to deal with the dynamic and complex environment of DTNs resulting from users mobility or varying conditions of wireless communications. CGrAnt is based on (1) Cultural Algorithms (CA) and Ant Colony Optimization (ACO) and (2) metrics which characterize opportunistic social connectivity between wireless users. CA and ACO are used to direct the network traffic, taking into account a set of social-aware metrics that may infer relevant structures in meeting regularities and mobility patterns of users. The most promising message forwarders are selected through a greedy transition rule based on local and global information captured from the DTN environment. Through simulation, we have analyzed the influence of ACO operators and CA's knowledge on CGrAnt performance. We have then compared the performance of CGrAnt with PROPHET and Epidemic protocols under varying networking parameters. Results have shown that CGrAnt achieves the highest delivery ratio and lowest byte redundancy.

6.7.1.3. Opportunistic Content Dissemination

Here, 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 related issues, 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. In this environment, opportunistic communication will not be effective due to the lack of known friends within the communication range. Thus, we have proposed a novel architecture that addresses the issues of lack of trust, timeliness of delivery, loss of user control, and privacy-aware distributed mobile social networking by combining the advantages of distributed decentralized storage and opportunistic communications. We have formally defined a content replication problem in mobile social networks and show that it is computationally hard to solve optimally. Then, we have proposed a community based greedy heuristic algorithm with novel dynamic centrality metrics to replicate content in well-selected users, to maximize the content dissemination with limited number of replication. Using both real world and synthetic traces, we have shown that content replication can attain a large coverage gain and reduce the content delivery latency.

6.7.1.4. Data Offloading-aware Hotspot Deployment

With the steady growth of sales of smart-phones, the demand for services that generate mobile data traffic has grown tremendously. The growing use of traffic data generated from mobile devices overloads the network infrastructure, which is not always prepared to receive such demand. To tackle this problem, we are studying the mobile behavior and resource consumptions of people on a metropolitan area in a major city and turn it into a set of well located WiFi hotspots. For this, we have proposed a data offloading-aware hotspot deployment. It

is methodologically divided as (i) creation of a time dependent weighted graph to represent people's mobility, traffic and its relation with places/locations able to receive a hotspot, (ii) measurement of location's importance and selection of the best-ranked ones. Better positioned hotspots are likely to provide better coverage, and therefore, be able to offload more data.

6.7.2. Collaborations

- Professors Anelise Munaretto and Myriam Regattieri Delgado from Federal Technological University of Parana (UTFPR), Brazil,
- Professors Aruna Seneviratne and Henrik Petander from NICTA and School of EE&T, UNSW, Sydney, Australia,
- Pedro O.S. Vaz de Melo and Antonio A. F. Loureiro, Federal University of Minas Gerais, Brazil,
- Marco Fiore and Frederic Le Mouel from INSA Lyon, France,
- Katia Jaffrès-Runser, University of Toulouse, IRIT/ENSEEIHT, France.

6.8. New services and protocols

Participants: Aline Carneiro Viana, Guilherme Maia.

6.8.1. Scientific achievements

6.8.1.1. Network Discovery

Network discovery is a fundamental task in different scenarios of IEEE 802.15.4-based wireless personal area networks. Scenario examples are body sensor networks requiring health- and wellness-related patient monitoring or situations requiring opportunistic message propagation. Therefore, we have investigated optimized discovery of IEEE 802.15.4 static and mobile networks operating in multiple frequency bands and with different beacon intervals. We designed a linear programming model that allows finding two optimized strategies, named OPT and SWOPT, to deal with the asynchronous and multi-channel discovery problem. We have also proposed a simplified discovery solution, named SUBOPT, featuring a low-complexity algorithm requiring less memory usage. A cross validation between analytical, simulation, and experimental evaluation methods was performed. Our performance studies confirmed improvements achieved by our solutions in terms of first, average, and last discovery time as well as discovery ratio, when compared to IEEE 802.15.4 standard approach and the SWEEP approach known from the literature.

6.8.1.2. Distributed Data Storage

The deployment of large-scale Wireless Sensor Network (WSN) applications (e.g., environment sensing and military surveillance), which operate unattended for long periods of time and generate a considerable amount of data, poses several challenges. One of them is *how to retrieve the sensed data*. To tackle this issue, we have designed ProFlex, a distributed data storage protocol for large-scale heterogeneous wireless sensor networks (HWSNs) with mobile sinks. ProFlex guarantees robustness in data collection by intelligently managing data replication among selected storage nodes in the network. Contrarily to related protocols in the literature, ProFlex considers the resource constraints of sensor nodes and constructs multiple data replication structures, which are managed by more powerful nodes. Additionally, ProFlex takes advantage of the higher communication range of such powerful nodes and uses the long-range links to improve data distribution by storage nodes. When compared with related protocols, we have shown through simulation that Proflex has an acceptable performance under message loss scenarios, decreases the overhead of transmitted messages, and decreases the occurrence of the energy hole problem. Moreover, we have proposed an improvement that allows the protocol to leverage the inherent data correlation and redundancy of wireless sensor networks in order to decrease even further the protocol's overhead without affecting the quality of the data distribution by storage nodes.

6.8.2. Collaborations

- PhD Niels Karowski, Technische Universitat Berlin, Germany,
- Professor Adam Wolisz, Technische Universitat Berlin, Germany,
- Antonio A. F. Loureiro, Federal University of Minas Gerais, Brazil,

RAP Project-Team

4. New Results

4.1. Algorithms: Bandwidth Allocation in Optical Networks

Participants: Christine Fricker, Philippe Robert, James Roberts.

The development of dynamic optical switching is widely recognized as an essential requirement to meet anticipated growth in Internet traffic. Since September 2009, RAP has investigated the traffic management and performance evaluation issues that are particular to this technology. A first analysis of passive optical networks used for high speed Internet access led to the proposal of an original dynamic bandwidth allocation algorithm and to an evaluation of its traffic capacity. Our activity on optical networking is carried out in collaboration with Orange Labs with whom we have a research contract. We have also established contacts with Alcatel-Lucent Bell Labs and had fruitful exchanges with Iraj Saniee and his team on their proposed time-domain wavelength interleaved networking architecture (TWIN).

We have analyzed the traffic capacity of wavelength division multiplexing (WDM), passive optical networks (PONs) where user stations (optical network units) are equipped with tunable transmitters. For these systems users can use any of the multiple wavelengths to transmit their data but only within the limit determined by the number of transmitters they possess. A mean field approximation is used to estimate the capacity of a limited-gated multiserver polling system with a limit on the number of servers a given station can use simultaneously. The approximation provides an expression for the stability limit under very general assumptions about the traffic process and system configuration.

In 2011, we began work on bandwidth allocation in meshed networks. We have evaluated the TWIN architecture in a metropolitan area network with an original medium access control (MAC) algorithm. This algorithm was inspired by our prior work on access networks and ensures an efficient and fair allocation of bandwidth to flows between network nodes.

The TWIN architecture is not extensible to a wide area for reasons of scalability and the excessive signalling delay between geographically distant nodes. We have therefore invented a new notion of a multipoint-to-multipoint lightpath that avoids these problems. A patent relating to this invention has been granted. This patent is owned by Orange following the terms of our contract with them. The paper [16] describes the invention and its evaluation. A major advantage demonstrated in this paper is the energy saving achieved by the use of the proposed optical technology in place of electronic routers. An extended version of the paper has been accepted for publication in Journal of Optical Communication and Networking [24].

Ongoing research seeks to apply this type of networking solution to data centres, on one hand, and to geographically spread tier-1 Internet carrier networks, on the other. Some of this work is performed in collaboration with Orange Labs under the terms of our research contract. An interesting new development is the application of new coherent optical technology that allows tunable receivers as well as tunable transmitters. We are evaluating the performance of a bandwidth allocation algorithm that exploits this technology.

A wider reaching collaboration has been established under the terms of a Celtic Plus project called SASER. This project was approved by the EU in 2012 and funding has been obtained for our participation from the French authorities. The project kickoff meeting was held in November 2012. Our contribution relates to the use of TWIN to create an extended metropolitan optical network. Our partners in the corresponding work package task are Orange, Telecom Bretagne and the engineering school ENSSAT. Overall reponsibility for the work package (where alternative optical network architectures are also evaluated) is with Alcatel-Lucent Bell Labs.

4.2. Algorithms: Content-Centric Networking

Participants: Mathieu Feuillet, Christine Fricker, Philippe Robert, James Roberts, Nada Sbihi.

RAP is participating in an ANR project named CONNECT which contributes to the definition and evaluation of a new paradigm for the future Internet: a content-centric network (CCN) where, rather than interconnecting remote hosts like IP, the network directly manages the information objects that users publish, retrieve and exchange. CCN has been proposed by Van Jacobson and colleagues at the Palo Alto Research Center (PARC). In CCN, content is divided into packet-size chunks identified by a unique name with a particular hierarchical structure. The name and content can be cryptographically encoded and signed, providing a range of security levels. Packets in CCN carry names rather than addresses and this has a fundamental impact on the way the network works. Security concerns are addressed at the content level, relaxing requirements on hosts and the network. Users no longer need a universally known address, greatly facilitating management of mobility and intermittent connectivity. Content is supplied under receiver control, limiting scope for denial of service attacks and similar abuse. Since chunks are self-certifying, they can be freely replicated, facilitating caching and bringing significant bandwidth economies. CCN applies to both stored content and to content that is dynamically generated, as in a telephone conversation, for example. RAP is contributing to the design of CCN in two main areas:

- the design and evaluation of traffic controls, recognizing that TCP is no longer applicable and queue management will require new, name-based criteria to ensure fairness and to realize service differentiation;
- the design and evaluation of replication and caching strategies that realize an optimal trade-off of expensive bandwidth for cheap memory.

The team also contributes to the development of efficient forwarding strategies and the elaboration of economic arguments that make CCN a viable replacement for IP. CONNECT partners are Alcatel-Lucent (lead), Orange, Inria/RAP, Inria/PLANETE, Telecom ParisTech, UPMC/LIP6.

A paper describing a proposed flow-aware approach for CCN traffic management and its performance evaluation has been presented at the conference Infocom 2012 [20]. We have reviewed the literature on cache performance (dating from early work on computer memory management) and identified a practical and versatile tool for evaluating the hit rate (proportion of requests that are satisfied from the cache) as a function of cache size and the assumed object popularity law. This approximate method was first proposed in 2002 by Che, Tung and Wang for their work on web caching. We applied this approximation to evaluate CCN caching performance taking into account the huge population and diverse popularity characteristics that make other approaches ineffective [19]. The excellent accuracy of this method over a wide range of practically relevant traffic models has been explained mathematically [18]. CONNECT ends in December 2012. We are currently defining a new project proposal that should be submitted to the ANR INFRA call in February 2013.

4.3. Scaling Methods: Fluid Limits in Wireless Networks

Participant: Philippe Robert.

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.

4.4. Algorithms: Distributed Hash Tables

Participants: Mathieu Feuillet, Philippe Robert.

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. We modeled this problem thanks to standard queues in order to understand the behavior of a single file and the global dynamic of the system. 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 have been able to study different time scales of the system with an averaging phenomenon. A paper has been submitted on this subject for the case where there are at most two copies of each file [25]. An article for the general case is in preparation. A more sophisticated distributed model with mean field techniques is under investigation.

On the side of this project, we notably studied the distribution of hitting times of the classical Ehrenfest and Engset models by using martingale techniques, furthermore their asymptotic behavior has been analyzed when the size of the system increases to infinity [11].

4.5. Stochastic Modeling of Biological Networks

Participants: Emanuele Leoncini, Philippe Robert.

This is a collaboration with Vincent Fromion from INRA Jouy en Josas, which started on October 2010.

The goal is to propose a mathematical model of the production of proteins in prokaryotes. Proteins are biochemical compounds that play a key role in almost all the cell functions and are crucial for cell survival and for life in general. In bacteria the protein production system has to be capable to produce abut 2500 different types of proteins in different proportions (from few dozens for the replication machinery up to 100000 for certain key metabolic enzymes). Bacteria uses more than the 85% of their resources to the protein production, making it the most relevant process in these organisms. Moreover this production system must meet two opposing problems: on one side it must provide a minimal quantity for each protein type in order to ensure the smooth-running of the cell, on the other side an “overproduction policy” for all the proteins is infeasible, since this would impact the global performance of the system and of the bacterium itself.

Gene expression is intrinsically a stochastic process: gene activation/deactivation occurs by means the encounter of polymerase/repressor with the specific gene, moreover many molecules that take part in the protein production act at extremely low concentrations. We have restated mathematically the classical model using Poisson point processes (PPP). This representation, well-known in the field of queueing networks but, as far as we know, new in the gene expression modeling, allowed us to weaken few hypothesis of the existing models, in particular the Poisson hypothesis, which is well-suited in some cases, but that, in some situations, is far from the biological reality as we consider for instance the protein assemblage. See [12].

The theoretical environment of Poisson point processes has lead us to propose a new model of gene expression which captures on one side the main mechanisms of the gene expression and on the other side it tries to consider hypothesis that are more significant from a biological viewpoint. In particular we have modeled: gene activation/deactivation, mRNA production and degradation, ribosome attachment on mRNA, protein elongation and degradation. We have shown how the probability distribution of the protein production and the protein lifetime may have a significant impact on the fluctuations of the number of proteins. We have obtained analytic formulas when the duration of protein assemblage and degradation follows a general probability distribution, i.e. without the Poisson hypothesis. In particular, by using a PPP representation we have been able to include the deterministic continuous phenomenon of protein degradation, which is the main protein degradation mechanism for stable proteins. We have showed moreover that this more realistic description is surprisingly identical in distribution with the classic assumption of protein degradation by means of a degrading protein (*proteosome*). We have used our model also to compare the variances resulting by choosing different hypotheses for the probability elongation, in particular we have hypothesize the protein assembly to be deterministic. This assumption is justified because of the elongation step, which consists of a large number

of elementary steps, can be described by the sum of exponential steps and the resulting distribution is well approximated by a Gaussian distribution because of the central limit theorem. Under the hypothesis of small variance of the resulting Gaussian distribution, we can assume the elongation step to be deterministic. The model has showed how, under the previous hypothesis, the variance on the number of proteins is bigger than the classical model with the Poisson hypothesis.

We have developed a C++ stochastic simulator for our general model, which has allowed the computation of variance when it was not possible to derive explicit analytic close formulas and the simulation of some extension of the actual model.

4.6. Stochastic Networks: Large Bike Sharing Systems

Participants: Christine Fricker, Hanene Mohamed, Danielle Tibi.

This is a collaboration with Nicolas Gast (EPFL) starting in December 2010. Bike sharing systems were launched by numerous cities as a part of urban transportation, for example Velib in 2007 (20 000 bikes, 1 500 stations). One of the major issues is the availability of the resources: bikes or free slots. These systems become a hot topic in Operation Research but studies on these stochastic networks are very few. To our knowledge, no theoretical study of such bike sharing systems exist taking into account the limited capacity of the stations.

We modeled this system in a symmetric case. Mean field limit gives the dynamic of a large system and the limiting stationary behavior of a single station as the system gets large. Analytical results are obtained and convergence proved in the standard model via Lyapounov functions. It allows to find the best ratio of bikes par station and to measure the improvement of incentive mechanisms, as choosing among two stations for example. Redistribution by trucks is also investigated. See [26].

Further results have been obtained for some heterogeneous systems. By mean field techniques, analytical results are obtained with Hanene Mohamed for systems with clusters (see [17]).

In a work in progress with *Danielle Tibi* , a more direct method is used when the network has a product form invariant measure by central and local limit theorem. It is a way to prove in this case the equivalence of ensembles, known in physic statistics. It applies to the simplest non homogeneous model. It gives a way to generalize the cluster case.

4.7. Random Graphs

Participant: Nicolas Broutin.

4.7.1. Connectivity in models of wireless networks

This is joint work with S. Boucheron (Paris 7), L. Devroye (McGill), N. Fraiman (McGill), and G. Lugosi (Pompeu Fabra).

The traditional models for wireless networks rely on geometric random graphs. However, if one wants to ensure that the graph be fully connected the radius of influence (hence the power necessary, and number of links) is too large to be fully scalable. Recently some models have been proposed that skim the neighbours and only retain a random subset for each node, hence creating a sparser overlay that would hopefully be more scalable. The first results on the size of the subsets which guarantee connectivity of overlay (the irrigation graph) [3] confirm that the average number of links per node is much smaller, but it remains large. These results motivate further investigations on the size of the largest connected component when one enforces a constant average degree which are in the process of being written.

4.7.2. Random graphs and minimum spanning trees

This is a long term collaboration with L. Addario-Berry (McGill), C. Goldschmidt (Oxford) and G. Miermont (ENS Lyon).

The random graph of Erdős and Rényi is one of the most studied models of random networks. Among the different ranges of density of edges, the "critical window" is the most interesting, both for its applications to the physics of phase transitions and its applications to combinatorial optimization (minimum spanning tree, constraint satisfaction problems). One of the major questions consists in determining the distribution of distances between the nodes. A limit object (a scaling limit) has been identified, that allows to describe precisely the first order asymptotics of pairwise distances between the nodes. This limit object is a random metric space whose definition allows to exhibit a strong connection between random graphs and the continuum random tree of Aldous. A variety of questions like the diameter, the size of cycles, etc, may be answered immediately by reading them on the limit metric space [2].

In a stochastic context, the minimum spanning tree is tightly connected to random graphs via Kruskal's algorithm. Random minimum spanning trees have attracted much research because of their importance in combinatorial optimization and statistical physics; however, until now, only parameters that can be grasped by local arguments had been studied. The scaling limit of the random graphs obtained in [2] permits to describe precisely the metric space scaling limit of a random minimum spanning tree [21], which identifies a novel continuum random tree which is truly different from that of Aldous.

4.7.3. Analysis of recursive partitions

This is joint work with R. Neininger (Frankfurt) and H. Sulzbach (Frankfurt/McGill).

The quadtrees are essential data structures that permit to store and manipulate geometric data by building a recursive partition of the space. In order to evaluate their performance, Flajolet and his co-authors have estimated the average cost of reporting all the data matching certain random queries. When the query does not fully specify all the fields, one talk about a partial match query. Such queries are ubiquitous, but analyzing their behaviour turns out to be intricate, and no performance guarantee was available in the form of a bound on the probability that any query would take much more time that one expects. [14] provides such guarantees by analysing the behaviour of all the queries at the same time, as a process. This yields estimates for the cost of the worst possible query (not a uniformly random one), as well as asymptotics for the variance and higher moments.

This line of research has motivated the analysis of the related combinatorial model of recursive lamination of the disk. The model had been recently introduced, but no full analysis was available. The techniques developed in the context of quadtrees have inspired a proof that the dual tree of the recursive lamination does converge to a limit tree-like metric space which is identified [23].

4.7.4. Navigation and point location in Poisson Delaunay triangulation

Nicolas Broutin has recently initiated a project with O. Devillers (Inria Sophia) and R. Hemsley (Inria Sophia) concerning the performance of local routing algorithms in plane subdivisions. Such algorithms also turn out to be important for the *point location* problem: for instance, finding the face of the subdivision which contains a query point is the first step towards inserting this point as a vertex. The aim is to prove that when the subdivision consists of the faces of a Delaunay triangulation, and when the points are random, any natural strategy which would take you closer to the aim performs well. Preliminary results about a specific routing algorithm, the cone walk, that we designed for its amenability to analysis appear in [22].

4.8. Stochastic Networks: Jackson Networks

Participant: Danielle Tibi.

Lyapounov functions and essential spectral radius of Jackson networks, joint work with I. Ignatiouk-Robert (University of Cergy-Pontoise). A family of explicit multiplicative Lyapounov functions is constructed for any stable Jackson network. Optimizing the multiplicative factor over this family provides an upper bound for the essential spectral radius of the associated Markov process. For some particular classes of Jackson networks, this upper bound coincides with a lower bound derived from large deviations arguments, thus providing the exact value of the essential spectral radius. The main example is given by Jackson networks with routing matrix having a tree structure (in the sense that for any node i, at most one other node can route its customers to i).

The result also holds for other types of routing matrices (e.g. completely symmetrical), under some conditions over the different arrival and service rates. See [27].

REGAL Project-Team

6. New Results

6.1. Introduction

In 2012, we focused our research on the following areas:

- *Management of distributed data.*
- *Performance and robustness of Systems Software in multicore architectures.*

6.2. Distributed algorithms for dynamic networks

Participants: Luciana Arantes [correspondent], Olivier Marin, Sébastien Monnet, Franck Petit [correspondent], Maria Potop-Butucaru, Pierre Sens, Julien Sopena, Raluca Diaconu, Ruijing Hu, Anissa Lamani, Sergey Legtchenko, Jonathan Lejeune, Karine Pires, Guthemberg Silvestre, Véronique Simon.

This objective aims to design distributed algorithms adapted to new large scale or dynamic distributed systems, such as mobile networks, sensor networks, P2P systems, Grids, Cloud environments, and robot networks. Efficiency in such demanding environments requires specialised protocols, providing features such as fault or heterogeneity tolerance, scalability, quality of service, and self-stabilization. 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.

Since 2008, 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.

6.2.1. Mutual Exclusion and Failure Detection.

Mutual Exclusion and Fault Tolerance are two major basic building blocks in the design of distributed systems. Most of the current mutual exclusion algorithms are not suitable for modern distributed architectures because they are not scalable, they ignore the network topology, and they do not consider application quality of service constraints. Under the ANR Project *MyCloud* and the FSE *Nu@age*, we study locking algorithms fulfilling some QoS constraints often found in Cloud Computing [38].

A classical way for a distributed system to tolerate failures is to detect them and then recover. It is now well recognized that the dominant factor in system unavailability lies in the failure detection phase. Regal has worked for many years on practical and theoretical aspects of failure detections and pioneered hierarchical scalable failure detectors. [2] Since 2008, we have studied the adaptation of failure detectors to dynamic networks. Following the model introduced in [18], we have proposed new algorithms to detect crashes and Byzantine behaviors [32].

These algorithms were designed as part of the ANR Project SHAMAN.

[2] Recent work by Leners et al published in SOSP 2011 uses our DSN 2003 paper as basis for performance comparison

6.2.2. Self-Stabilization and Self-* Services.

We have also approached fault tolerance through self-stabilization. Self-stabilization is a versatile technique to design distributed algorithms that withstand transient faults. In particular, we have worked on the unison problem, [3] i.e., the design of self-stabilizing algorithms to synchronize a distributed clock. As part of the ANR project *SPADES*, we have proposed several snap-stabilizing algorithms for the message forwarding problem that are optimal in terms of number of required buffers [36]. A snap-stabilizing algorithm is a self-stabilizing algorithm that stabilizes in 0 steps; in other words, such an algorithm always behaves according to its specification.

Finally, we have applied our expertise in distributed algorithms for dynamic and self-* systems in domains that at first glance seem quite far from the core expertise of the team, namely ad-hoc systems and swarms of mobile robots. In the latter, as part of ANR project *R-Discover*, we have studied various problems such as exploration [29], and gathering [15].

6.2.3. Dissemination and Data Finding in Large Scale Systems.

In the area of large-scale P2P networks, we have studied the problems of data dissemination and overlay maintenance, i.e., maintenance of a logical network built over the a P2P network. First, we have proposed efficient distributed algorithms to ensure data dissemination to a large set of nodes. Also, we have introduced a new method to compare dissemination algorithms over various topologies [35].

6.2.4. MMOGs.

Peer-to-peer overlay networks can be used to build scalable infrastructures for MMOGs. Our work on MMOGs has primarily focused on the impact of latency constraints in dynamic distributed systems. In online P2P games, players are connected by a logical graph, implemented as an overlay network. Latency constraints imply that players that interact must remain close in the overlay, even when the mobility of players induces rapid changes in the graph.

We have also addressed problems related to cheating and arbitration. In a distributed system, certification of entities makes it possible to circumscribe malicious behavior, such as cheating in games. Certification requires the use of a trusted third party and is traditionally done centrally. At a large scale, however, centralized certification represents a bottleneck and a single point of attack or failure. We have proposed solutions based on distributed reputations to identify trusted nodes and use them as game referees to detect and prevent cheating [46]. Our method relies on previous work on the subject of trusted node collaboration to ensure reliable distributed certification [4].

6.3. Management of distributed data

Participants: Mesaac Makpangou, Olivier Marin, Sébastien Monnet, Pierre Sens, Marc Shapiro, Julien Sopena, Gaël Thomas, Pierpaolo Cincilla, Raluca Diaconu, Sergey Legtchenko, Jonathan Lejeune, Karine Pires, Thomas Preud homme, Masoud Saeida Ardekani, Guthemberg Silvestre, Pierre Sutra, Marek Zawirski, Annette Bieniusa, Pierpaolo Cincilla, Véronique Simon, Mathieu Valero.

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 to ensure consistency between replicas; when and how to move data to computation, or computation to data, in order to improve response time while minimizing storage or energy usage; etc. The Regal group works on several key issues related to replication:

[3] C. Boulinier, F. Petit, and V. Villain. Synchronous vs. asynchronous unison. Algorithmica, 51(1):61-80, 2008

[4] Erika Rosas, Olivier Marin and Xavier Bonnaire. CORPS: Building a Community Of Reputable PeerS in Distributed Hash Tables. The Computer Journal, 54(10):1721-1735(2011)

- Replica placement for fault tolerance and latency in the presence of churn,
- scalable strong consistency for replicated databases, and
- theory and practice of eventual consistency.

6.3.1. Distributed hash tables

A DHTs replicates data and spreads the replicas uniformly across a large number of nodes. Being very scalable and fault-tolerant, DHTs are a key component for dependable and secure applications, such as backup systems, distributed file systems, multi-range query systems, and content distribution systems.

Despite the advantages of DHTs, several studies show that they become inefficient in environments subject to churn, i.e., with many node arrivals and departures. We therefore propose a new replication mechanism for DHTs that is churn resilient [20]. RelaxDHT relaxes placement constraints, in order to avoid redundant data transfers and to increase parallelism. RelaxDHT loses up to 50% fewer data blocks that the well-known PAST DHT.

6.3.2. 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.

To reduce the latency of consensus, we study *Generalised Consensus* algorithms, i.e., ones that leverage the commutativity of operations or the spontaneous ordering of messages by the network. We propose a novel protocol for generalised consensus that is optimal, both in message complexity and in faults tolerated, and that switches optimally between its fast path (which avoids ordering commuting requests) and its classical path (which generates a total order). Experimental evaluation shows that our algorithm is much more efficient and scales better than competing protocols.

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, namely that a system cannot have both genuine partial replication and SI. We designed an efficient protocol that maintains the most important features of SI, but side-steps this impossibility. Finally, we study the trade-offs between freshness (and hence low abort rates) and space complexity in computing snapshots, as required by SI and its variants.

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. It 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. Our simulations 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. This is published at ICPADS 2012 [27].

Our current experiments aim to compare the practical pros and cons of different approaches to designing large-scale replicated databases, by implementing and benchmarking a number of different protocols.

Our study the trade-offs between freshness and meta-date overhead, is published in HotCDP 2012 [43].

6.3.3. 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). In particular, we publish a study of the concurrency semantics of sets in DISC 2012 [48], [22].

On the theoretical side, we identified sufficient correctness conditions for CRDTs, viz., that concurrent updates commute, or that the state is a monotonic semi-lattice. 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?

We are currently developing a very large-scale CRDT platform called SwiftCloud, which aims to scale to millions of clients, deployed inside and outside the cloud.

6.4. Improving the Performance and Robustness of Systems Software in Multicore Architectures

6.4.1. Managed Runtime Environments

Participants: Bertil Folliot, Julia Lawall, Gilles Muller [correspondent], Marc Shapiro, Julien Sopena, Gaël Thomas, Florian David, Lokesh Gidra, Jean-Pierre Lozi, Thomas Preud homme, Suman Saha, Harris Bakiras, Arie Middelkoop, Koutheir Attouchi.

Today, multicore architectures are becoming ubiquitous, found even in embedded systems, and thus it is essential that managed languages can scale on multicore processors. We have found that a major scalability bottleneck is the implementation of high contention locks, which can overload the bus, eliminating all performance benefits from adding more cores. To address this issue, as part of the PhD of Jean-Pierre Lozi, we have developed remote core locking (RCL), in which highly contended locks are implemented on a dedicated server, minimizing bus traffic and improving application scalability (USENIX ATC 2012 [24]). This work initially targeted C code but is now being adapted to the needs of Java applications in the PhD of Florian David. Another bottleneck in the support for managed languages is the garbage collector. As part of the PhD of Lokesh Gidra, we have identified the main sources of overhead.

6.4.2. Systems software robustness

A new area of research for Regal, with the arrival of Gilles Muller in 2009 as Inria Senior Research Scientist and Julia Lawall in 2011 as Inria Senior Research Scientist, is on improving the reliability of operating systems code. Muller and Lawall previously developed Coccinelle, a scriptable program matching and transformation tool for C code that is now commonly used in the open-source development community, including by the developers of Linux, Wine and Dragonfly BSD. Based on Coccinelle, we have developed a new approach to inferring API function usage protocols from software, relying on knowledge of common code structures (Software – Practice and Experience [19]).

We have also proposed a method for automatically identifying bug-fixing patches, with the goal of helping developers maintain stable versions of the software (ICSE 2012 [45]) and have designed an approach to automatically generating a robust interface to the Linux kernel, to provide developers of new kernel-level code more feedback in the case of a misunderstanding of kernel API usage conventions (ASE 2012 [24]).

TREC Project-Team

6. New Results

6.1. Design and Performance Analysis of Wireless Networks

Participants: François Baccelli, Bartłomiej Błaszczyszyn, Chung Shue Chen, Miodrag Jovanović, Holger Paul Keeler, Mir Omid Haji Mirsadeghi, Frédéric Morlot, Tien Viet Nguyen.

CDMA/UMTS, Wireless LANs, ad hoc networks, IEEE 802.11, mesh networks, cognitive radio, Hiperlan, CSMA, TCP, MAC protocols, exponential back-off protocols, signal to interference ratio, coverage, capacity, transport capacity, admission and congestion control.

This axis bears on the analysis and the design of wireless access communication networks. Our contributions are organized in terms of network classes: cellular networks, wireless LANs and MANETs, VANETs. We also have a section on generic results that regard more general wireless networks. We are interested both in macroscopic models, which are particularly important for economic planning and in models allowing the definition and the optimization of protocols. Our approach combines several tools, queueing theory, point processes, stochastic geometry, random graphs, distributed control algorithms, self organization protocols.

6.1.1. Cellular Networks

The activity on cellular networks has several complementary facets ranging from performance evaluation to protocol design. The work is mainly based on strong collaborations with Alcatel-Lucent and Orange Labs.

6.1.1.1. Effect of Opportunistic Scheduling on the Quality of Service Perceived by the Users in OFDMA Cellular Networks

Our objective in [17] is to analyze the impact of fading and opportunistic scheduling on the quality of service perceived by the users in an Orthogonal Frequency Division Multiple Access (OFDMA) cellular network. To this end, assuming Markovian arrivals and departures of customers that transmit some given data volumes, as well as some temporal channel variability (fading), we study the mean throughput that the network offers to users in the long run of the system. Explicit formulas are obtained in the case of allocation policies, which may or may-not take advantage of the fading, called respectively opportunistic and non-opportunistic. The main practical results of the present work are the following. Firstly we evaluate for the non-opportunist allocation the degradation due to fading compared to Additive White Gaussian Noise (AWGN) (that is, a decrease of at least 13% of the throughput). Secondly, we evaluate the gain induced by the opportunistic allocation. In particular, when the traffic demand per cell exceeds some value (about 2 Mbits/s in our numerical example), the gain induced by opportunism compensates the degradation induced by fading compared to AWGN. Partial results were presented at ComNet in 2009 [61].

6.1.1.2. Impact of propagation-loss model on the geometry and performance of cellular networks

6.1.1.2.1. Impact of Shadowing on QoS

Shadowing is believed to degrade the quality of service in wireless cellular networks. In [18] we discovered a more subtle reality. Increasing variance of the lognormal shadowing tends to “separate” the strongest (serving BS) signal from all other signals — a phenomenon observed for heavy-tailed distributions and called “single big jump principle”. In consequence, in some cases, an increase of the variance of the shadowing can significantly reduce the mean interference factor and improve some QoS metrics in interference limited systems. We exemplify this phenomenon, similar to stochastic resonance and related to the “single big jump principle” of the heavy-tailed log-nornal distribution, studying the blocking probability in regular, hexagonal networks in a semi-analytic manner, using a spatial version of the Erlang’s loss formula combined with Kaufman-Roberts algorithm.

6.1.1.2.2. Using Poisson processes to model lattice cellular networks

In [51] we mathematically proved that a large spatially homogeneous (arbitrary, including hexagonal) network is perceived by a typical user as an equivalent (infinite) Poisson network, provided shadowing is strong enough. This justifies an almost ubiquitous Poisson assumption made in the stochastic-analytic approach to study of the quality of user-service in cellular networks.

6.1.1.2.3. Linear-Regression Estimation of the Propagation-Loss Parameters Using Mobiles' Measurements

In [35] we proposed a new linear-regression model for the estimation of the path-loss exponent and the parameters of the shadowing from the propagation-loss data collected by the mobiles with respect to their serving base stations. The model is based on the aforementioned Poisson convergence result.

6.1.1.3. Quality of Real-Time Streaming in Wireless Cellular Networks

In [50] we present a new stochastic service model with service capacity sharing and interruptions, meant to be useful for the performance evaluation and dimensioning of wireless cellular networks offering real-time streaming, like e.g. mobile TV. Our general model takes into account Markovian, multi-class process of call arrivals, arbitrary streaming time distribution, and allows for a general service (outage) policy saying which users are temporarily denied the service due to insufficient service capacity. Using Palm theory formalism, we develop expressions for several important characteristics of this model, including mean time spent in outage and mean number of outage incidents for a typical user of a given class. We also propose some natural class of least-effort-served-first service policies, for which the aforementioned expressions can be efficiently evaluated on the basis of the Fourier analysis of Poisson process. Last but not least, we show how our model can be used to analyse the quality of real-time streaming in 3GPP Long Term Evolution (LTE) cellular networks. We identify and evaluate an optimal and a fair service policy, the latter being suggested by LTE implementations, as well as propose some intermediate policies which allow to solve the optimality/fairness trade-off caused by unequal user radio-channel conditions.

6.1.1.4. Theoretically Feasible QoS in a MIMO Cellular Network Compared to the Practical LTE Performance

Our goal in [39] is to build a global analytical approach for the evaluation of the quality of service perceived by the users in wireless cellular networks which is calibrated in some reference cases. To do so, a model accounting for interference in a MIMO cellular system is firstly described. An explicit expression of users bit-rates theoretically feasible from the information theory point of view is then deduced. The comparison between these bit-rates and practical LTE performance permits to obtain the progress margins for potential evolution of the technology. Moreover, it leads to an analytical approximate expression of the system performance which is calibrated with the practical one. This expression is the keystone of a global analytical approach for the evaluation of the QoS perceived by the users in the long run of users arrivals and departures in the network. We illustrate our approach by calculating the users QoS as function of the cell radius in different mobility and interference cancellation scenarios.

6.1.1.5. Self-Optimization of Radio Resources in Cellular Networks

In [19], we surveyed the mathematical and algorithmic tools for the self-optimization of mobile cellular networks based on Gibbs' sampler. This technique allows for the joint optimization of radio resources in heterogeneous cellular networks made of a juxtaposition of macro and small cells. It can be implemented in a distributed way and nevertheless achieves minimal system-wide potential delay. Results show that it is effective in both throughput and energy efficiency.

Three patents were filed on this line of thought under the Inria/Alcatel–Lucent joint laboratory.

6.1.1.6. Coverage in Cellular Networks

Cellular networks are in a major transition from a carefully planned set of large tower-mounted base-stations (BSs) to an irregular deployment of heterogeneous infrastructure elements that often additionally includes micro, pico, and femtocells, as well as distributed antennas. In a collaboration with H. Dhillon, J. Andrews and R. Ganti [UT Austin, USA] [20], we developed a model for a downlink heterogeneous cellular network (HCN) consisting of K tiers of randomly located BSs, where each tier may differ in terms of average transmit power, supported data rate and BS density. Assuming a mobile user connects to the strongest candidate BS, the

resulting Signal-to-Interference-plus-Noise-Ratio (SINR) is greater than 1 when in coverage, Rayleigh fading, we derived an expression for the probability of coverage (equivalently outage) over the entire network under both open and closed access. One interesting observation for interference-limited open access networks is that at a given SINR, adding more tiers and/or BSs neither increases nor decreases the probability of coverage or outage when all the tiers have the same SINR threshold.

6.1.2. Mobile Ad Hoc Networks

A MANET is made of mobile nodes which are at the same time terminals and routers, connected by wireless links, the union of which forms an arbitrary topology. The nodes are free to move randomly and organize themselves arbitrarily. Important issues in such a scenario are connectivity, medium access (MAC), routing and stability. This year, we worked on a game theoretic view of Spatial Aloha in collaboration with E. Altman and M.K. Hanawal [Inria MAESTRO] [22] This line of though is currently continued with Chandramani Singh. We also compared the performance of spatial Aloha to CSMA.

6.1.2.1. Improvement of CSMA/CA's Spatial Reuse

The most popular medium access mechanism for such ad hoc networks is CSMA/CA with RTS/CTS. In CSMA-like mechanisms, spatial reuse is achieved by implementing energy based guard zones. In a collaboration with Qualcomm [12], we considered the problem of simultaneously scheduling the maximum number of links that can achieve a given signal to interference ratio (SIR). Using tools from stochastic geometry, we studied and maximized the medium access probability of a typical link. Our contributions are two-fold: (i) We showed that a simple modification to the RTS/CTS mechanism, viz., changing the receiver yield decision from an energy-level guard zone to an SIR guard zone, leads to performance gains; and (ii) We showed that this combined with a simple modification to the transmit power level – setting it to be inversely proportional to the square root of the link gain – leads to significant improvements in network throughput. Further, this simple power-level choice is no worse than a factor of two away from optimal over the class of all "local" power level selection strategies for fading channels, and further is optimal in the non-fading case. The analysis relies on an extension of the Matérn hard core point process which allows us to quantify both these SIR guard zones and this power control mechanism.

6.1.2.2. Comparison of the maximal spatial throughput of Aloha and CSMA in Wireless multihop Ad-Hoc Networks

In [46] this paper we compare the spatial throughput of Aloha and Carrier Sense Multiple Access (CSMA) in Wireless multihop Ad-Hoc Networks. In other words we evaluate the gain offered by carrier sensing (CSMA) over the pure statiscal collision avoidance which is the basis of Aloha. We use a Signal-to-Interference-and-Noise Ratio (SINR) model where a transmission is assumed to be successful when the SINR is larger than a given threshold. Regarding channel conditions, we consider both standard Rayleigh and negligible fading. For slotted and non-slotted Aloha, we use analytical models as well as simulations to study the density of successful transmissions in the network. As it is very difficult to build precise models for CSMA, we use only simulations to compute the performances of this protocol. We compare the two Aloha versions and CSMA on a fair basis, i.e. when they are optimized to maximize the density of successful transmissions. For slotted Aloha, the key optimization parameter is the medium access probability, for non-slotted Aloha we tune the mean back-off time, whereas for CSMA it is the carrier sense threshold that is adjusted. Our study shows that CSMA always outperforms slotted Aloha, which in turn outperforms its non-slotted version.

6.1.2.3. Stochastic Analytic Evaluation of End-to-End Performance of Linear Nearest Neighbour Routing in MANETs with Aloha

Planar Poisson models with the Aloha medium access scheme have already proved to be very useful in studies of mobile ad-hoc networks (MANETs). However, it seems difficult to quantitatively study the performances of end-to-end routing in these models. In order to tackle this problem, in [52], we study a linear stationary route embedded in an independent planar field of interfering nodes. We consider this route as an idealization of a "typical" route in a MANET obtained by some routing mechanism. Such a decoupling allows us to obtain many numerically tractable expressions for local and mean end-to-end delays and the speed of packet progression, assuming slotted Aloha MAC and the Signal-to-Interference-and-Noise Ratio (SINR) capture condition, with the usual power-law path loss model and Rayleigh fading. These expressions show how the

network performance depends on the tuning of Aloha and routing parameters and on the external noise level. In particular we show a need for a well-tuned lattice structure of fixed relaying nodes, which helps to relay packets on long random routes in the presence of a non-negligible noise. We also consider a Poisson-line MANET model, in which nodes are located on roads forming a Poisson-line process. In this case our linear route is rigorously (in the sense of Palm theory) the typical route in this Poisson-line MANET.

6.1.3. Vehicular Ad-Hoc Networks (VANETs)

Vehicular Ad Hoc NETworks (VANETs) are special cases of MANETs where the network is formed between vehicles. VANETs are today the most promising civilian application for MANETs and they are likely to revolutionize our traveling habits by increasing safety on the road while providing value added services.

6.1.3.1. Point-to-Point, Emergency and Broadcast Communications

Our aim in [36] is to analyze the Aloha medium access (MAC) scheme in one-dimensional, linear networks, which might be an appropriate assumption for VANETs. The locations of the vehicles are assumed to follow a homegeneous Poisson point process. Assuming powerlaw mean path-loss and independent point-to-point fading we study performance metrics based on the signal-over-interference and noise ratio (SINR). In contrast to previous studies where the receivers are at a fixed distance from the transmitter, we assume here that the receivers are the nearest neighbors of the transmitters in the Poisson process and in a given direction. We derive closed formulas for the capture probability and for the density of progress of a packet sent by a given node. We compute the mean delay to send a packet transmitted at each slot until successful reception. We also evaluate an upper bound to discover the neighborhood within a given space interval. We show that we can include noise in the previous models.

6.1.4. Cognitive Radio Networks

We wrote a survey [26] on the probabilistic framework which can be used to model and analyze cognitive radio networks using various classes of MAC protocols (including carrier sensing based multiple access schemes and Aloha schemes). For each model, analytical results were derived for important performance metrics. This leads to a quantification of the interplay between primary and secondary users in such networks.

6.2. Network Dynamics

Participants: Abir Benabid, Julieta Bollati, Anne Bouillard, Ana Bušić, Emilie Coupechoux, Nadir Farhi.

Queueing network, stability, inversion formula, probing, estimator, product-form, insensitivity, markov decision, max-plus algebra, network calculus.

6.2.1. Network Calculus

Network calculus is a theory that aims at computing deterministic performance guarantees in communication networks. This theory is based on the (min,plus) algebra. Flows are modeled by an *arrival curve* that upper-bounds the amount of data that can arrive during any interval, and network elements are modeled by a *service curve* that gives a lower bound on the amount of service offered to the flows crossing that element. Worst-case performances are then derived by combining these curves.

6.2.1.1. Performance bounds in FIFO tandem networks

In cooperation with Giovanni Stea [University of Pisa, Italy], we present in [31] algorithms to compute worst-case performance upper bounds when the service policy is FIFO, using linear programming. Linear programming leads to tight bounds; however, the computation corst is too high for reasonable-size networks. We then develop approximate solution schemes to find both upper and lower delay bounds on the worst-case delay. Both of them only require to solve just one LP problem, and they produce bounds which are generally more accurate than those found in the literature. Finally, we have a conjecture on what sould be the worst-case trajectory under usual assumptions.

6.2.1.2. Feed-forward networks with wormhole routing discipline

In collaboration with Bruno Gaujal [Inria Rhone Alpes] and Nadir Farhi [IFFSTAR] we are working on a model of performance bound calculus on feed-forward networks where data packets are routed under wormhole routing discipline. We are interested in determining maximum end-to-end delays and backlogs for packets going from a source node to a destination node, through a given virtual path in the network. Our objective is to give a "network calculus" approach to calculate the performance bounds. For this, we propose a new concept of curves that we call *packet curves*. The curves permit to model constraints on packet lengths for data flows, when the lengths are allowed to be different. We used this new concept to propose an approach for calculating residual services for data flows served under non preemptive service disciplines. This notion also enabled us to differentiate different classes of service policies: those that are based on a packet count (like round-robin and its generalized version), where the packet curve will be useful to tighten the bounds computed, and those that are based on the amount of data served (FIFO, priorities), where it won't be useful. These results have been presented at Valuetools (invited paper, [29]).

6.2.1.3. Using arrival curves for detecting anomalies in a network

In cooperation with Aurore Junier [Inria/IRISA] and Benoît Ronot [Alcatel-Lucent], we present an on-line algorithm that performs a flow of messages analysis. More precisely, it is able to highlight hidden abnormal behaviors that existing network management methods would not detect. Our algorithm uses the notion of constraint curves, introduced in the Network Calculus theory, defining successive time windows that bound the flow. The advantage of this algorithm is that it can be performed online, and in a second version has different levels of precision. This work has been presented in [30] and a patent [57] has been submitted.

6.2.1.4. Min,plus algorithms for fast weak-KAM integrators

In cooperation with Erwan Faou [IPSO-Inria Rennes, DMA-ENS] and Maxime Zavidovique [Paris 6]. We consider a numerical scheme for Hamilton-Jacobi equations based on a direct discretization of the Lax-Oleinik semi-group. We prove that this method is convergent with respect to the time and space stepsizes provided the solution is Lipschitz, and give an error estimate. Moreover, we prove that the numerical scheme is a geometric integrator satisfying a discrete weak-KAM theorem which allows to control its long time behavior. Taking advantage of a fast algorithm for computing min-plus convolutions based on the decomposition of the function into concave and convex parts, we show that the numerical scheme can be implemented in a very efficient way. The results can be found in [49].

6.2.2. Perfect Sampling of Queueing Systems

Propp and Wilson introduced in 1996 a perfect sampling algorithm that uses coupling arguments to give an unbiased sample from the stationary distribution of a Markov chain on a finite state space $\mathcal{X}$. In the general case, the algorithm starts trajectories from all $x \in \mathcal{X}$ at some time in the past until time $t = 0$. If the final state is the same for all trajectories, then the chain has coupled and the final state has the stationary distribution of the Markov chain. Otherwise, the simulations are started further in the past. This technique is very efficient if all the events in the system have appropriate monotonicity properties. However, in the general (non-monotone) case, this technique requires that one consider the whole state space, which limits its application only to chains with a state space of small cardinality.

6.2.2.1. Piecewise Homogeneous Events

In collaboration with Bruno Gaujal [Inria Grenoble - Rhone-Alpes], we proposed in [15] a new approach for the general case that only needs to consider two trajectories. Instead of the original chain, we used two bounding processes (envelopes) and we showed that, whenever they couple, one obtains a sample under the stationary distribution of the original chain. We showed that this new approach is particularly effective when the state space can be partitioned into pieces where envelopes can be easily computed. We further showed that most Markovian queueing networks have this property and we propose efficient algorithms for some of them.

The envelope technique has been implemented in a software tool PSI2 (see Section 5.2).

6.2.2.2. Perfect Sampling of Networks with Finite and Infinite Capacity Queues

In [33], we consider open Jackson queueing networks with mixed finite and infinite buffers and analyze the efficiency of sampling from their exact stationary distribution. We show that perfect sampling is possible, although the underlying Markov chain has a large or even infinite state space. The main idea is to use a Jackson network with infinite buffers (that has a product form stationary distribution) to bound the number of initial conditions to be considered in the coupling from the past scheme. We also provide bounds on the sampling time of this new perfect sampling algorithm under hyper-stability conditions (to be defined in the paper) for each queue. These bounds show that the new algorithm is considerably more efficient than existing perfect samplers even in the case where all queues are finite. We illustrate this efficiency through numerical experiments.

6.2.3. Markov Chains and Markov Decision Processes

Solving Markov chains is in general difficult if the state space of the chain is very large (or infinite) and lacking a simple repeating structure. One alternative to solving such chains is to construct models that are simple to analyze and provide bounds for a reward function of interest. The bounds can be established by using different qualitative properties, such as stochastic monotonicity, convexity, submodularity, etc. In the case of Markov decision processes, similar properties can be used to show that the optimal policy has some desired structure (e.g. the critical level policies).

6.2.3.1. Stochastic Monotonicity

In collaboration with Jean-Michel Fourneau [PRiSM, Université de Versailles Saint-Quentin] we consider two different applications of stochastic monotonicity in performance evaluation of networks [14]. In the first one, we assume that a Markov chain of the model depends on a parameter that can be estimated only up to a certain level and we have only an interval that contains the exact value of the parameter. Instead of taking an approximated value for the unknown parameter, we show how we can use the monotonicity properties of the Markov chain to take into account the error bound from the measurements. In the second application, we consider a well known approximation method: the decomposition into submodels. In such an approach, models of complex networks are decomposed into submodels whose results are then used as parameters for the next submodel in an iterative computation. One obtains a fixed point system which is solved numerically. In general, we have neither an existence proof of the solution of the fixed point system nor a convergence proof of the iterative algorithm. Here we show how stochastic monotonicity can be used to answer these questions. Furthermore, monotonicity properties can also help to derive more efficient algorithms to solve fixed point systems.

6.2.3.2. Markov Reward Processes and Aggregation

In a joint work with I.M. H. Vliegen [University of Twente, The Netherlands] and A. Scheller-Wolf [Carnegie Mellon University, USA] [16], we presented a new bounding method for Markov chains inspired by Markov reward theory: Our method constructs bounds by redirecting selected sets of transitions, facilitating an intuitive interpretation of the modifications of the original system. We show that our method is compatible with strong aggregation of Markov chains; thus we can obtain bounds for an initial chain by analyzing a much smaller chain. We illustrated our method by using it to prove monotonicity results and bounds for assemble-to-order systems.

6.2.3.3. Bounded State Space Truncation

Markov chain modeling often suffers from the curse of dimensionality problems and many approximation schemes have been proposed in the literature that include state-space truncation. Estimating the accuracy of such methods is difficult and the resulting approximations can be far from the exact solution. Censored Markov chains (CMC) allow to represent the conditional behavior of a system within a subset of observed states and provide a theoretical framework to study state-space truncation. However, the transition matrix of a CMC is in general hard to compute. Dayar et al. (2006) proposed DPY algorithm, that computes a stochastic bound for a CMC, using only partial knowledge of the original chain. In [32], we prove that DPY is optimal for the information they take into account. We also show how some additional knowledge on the chain can improve stochastic bounds for CMC.

6.2.4. Dynamic Systems with Local Interactions

Dynamic systems with local interactions can be used to model problems in distributed computing: gathering a global information by exchanging only local information. The challenge is two-fold: first, it is impossible to centralize the information (cells are indistinguishable); second, the cells contain only a limited information (represented by a finite alphabet $\mathcal{A}$; $\mathcal{A} = \{0, 1\}$ in our case). Two natural instantiations of dynamical systems are considered, one with synchronous updates of the cells, and one with asynchronous updates. In the first case, time is discrete, all cells are updated at each time step, and the model is known as a *Probabilistic Cellular Automaton (PCA)* (e.g. Dobrushin, R., Kryukov, V., Toom, A.: *Stochastic cellular systems: ergodicity, memory, morphogenesis*, 1990). In the second case, time is continuous, cells are updated at random instants, at most one cell is updated at any given time, and the model is known as a (finite range) *Interacting Particle System (IPS)* (e.g. Liggett, T.M.: *Interacting particle systems*, 2005).

6.2.4.1. Density Classification on Infinite Lattices and Trees

In a joint work with N. Fatès [Inria Nancy – Grand-Est], J. Mairesse and I. Marcovici [LIAFA, CNRS and Université Paris 7] [43] we consider an infinite graph with nodes initially labeled by independent Bernoulli random variables of parameter p. We address the density classification problem, that is, we want to design a (probabilistic or deterministic) cellular automaton or a finite-range interacting particle system that evolves on this graph and decides whether p is smaller or larger than $1/2$. Precisely, the trajectories should converge (weakly) to the uniform configuration with only $0's$ if $p < 1/2$, and only $1's$ if $p > 1/2$. We present solutions to that problem on $\mathbb{Z}^d$, for any $d \geq 2$, and on the regular infinite trees. For $\mathbb{Z}$, we propose some candidates that we back up with numerical simulations.

6.3. Economics of Networks

Participants: François Baccelli, Emilie Coupechoux, Marc Lelarge.

6.3.1. Diffusion and Cascading Behavior in Random Networks

The spread of new ideas, behaviors or technologies has been extensively studied using epidemic models. In [25], we consider a model of diffusion where the individuals' behavior is the result of a strategic choice. We study a simple coordination game with binary choice and give a condition for a new action to become widespread in a random network. We also analyze the possible equilibria of this game and identify conditions for the coexistence of both strategies in large connected sets. Finally we look at how can firms use social networks to promote their goals with limited information.

Our results differ strongly from the one derived with epidemic models. In particular, we show that connectivity plays an ambiguous role: while it allows the diffusion to spread, when the network is highly connected, the diffusion is also limited by high-degree nodes which are very stable. In the case of a sparse random network of interacting agents, we compute the contagion threshold for a general diffusion model and show the existence of (continuous and discontinuous) phase transitions. We also compute the minimal size of a seed of new adopters in order to trigger a global cascade if these new adopters can only be sampled without any information on the graph. We show that this minimal size has a non-trivial behavior as a function of the connectivity. Our analysis extends methods developed in the random graphs literature based on the properties of empirical distributions of independent random variables, and leads to simple proofs.

6.3.2. Coordination in Network Security Games: a Monotone Comparative Statics Approach

Malicious softwares or malwares for short have become a major security threat. While originating in criminal behavior, their impact are also influenced by the decisions of legitimate end users. Getting agents in the Internet, and in networks in general, to invest in and deploy security features and protocols is a challenge, in particular because of economic reasons arising from the presence of network externalities. In [24], [42], we focus on the question of incentive alignment for agents of a large network towards a better security. We start with an economic model for a single agent, that determines the optimal amount to invest in protection. The model takes into account the vulnerability of the agent to a security breach and the potential loss if a security breach occurs. We derive conditions on the quality of the protection to ensure that the optimal amount spent

on security is an increasing function of the agent's vulnerability and potential loss. We also show that for a large class of risks, only a small fraction of the expected loss should be invested. Building on these results, we study a network of interconnected agents subject to epidemic risks. We derive conditions to ensure that the incentives of all agents are aligned towards a better security. When agents are strategic, we show that security investments are always socially inefficient due to the network externalities. Moreover alignment of incentives typically implies a coordination problem, leading to an equilibrium with a very high price of anarchy.

6.4. Point Processes, Stochastic Geometry and Random Geometric Graphs

Participants: François Baccelli, Bartłomiej Błaszczyszyn, Pierre Brémaud, Kumar Gaurav, Mir Omid Haji Mirsadeghi.

stochastic geometry, point process, shot-noise, Boolean model, random tessellation, percolation, stochastic comparison

6.4.1. Modeling, comparison and impact of spatial irregularity of point processes on coverage, percolation, and other characteristics of random geometric models

We develop a general approach for comparison of clustering properties of point processes. It is funded on some basic observations allowing to consider void probabilities and moment measures as two complementary tools for capturing clustering phenomena in point processes. As 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 the void probabilities and moment measures, thus allowing to study the impact of clustering of the underlying point process. When stronger tools are needed, dcx 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 concerned. The whole approach has been worked out in a series of papers [62], [63], [64], [65]. This year we have prepared revisions of the two latter ones, from which [65] is now accepted for the publication in Adv. Appl. Probab. We have also prepared a review article [53] for *Lecture Notes in Mathematics*, Springer.

6.4.1.1. AB random geometric graphs

We investigated percolation in the AB Poisson-Boolean model in d-dimensional Euclidean space, and asymptotic properties of AB random geometric graphs on Poisson points in $[0,1]^d$. The AB random geometric graph we studied is a generalization to the continuum of a bi-partite graph called the AB percolation model on discrete lattices. Such an extension is motivated by applications to secure communication networks and frequency division duplex networks. The AB Poisson Boolean model is defined as a bi-partite graph on two independent Poisson point processes of intensities λ and μ in the d-dimensional Euclidean space in the same manner as the usual Boolean model with a radius r. We showed existence of AB percolation for all $d \geq 2$, and derived bounds for a critical intensity. Further, in $d = 2$, we characterize a critical intensity. The set-up for AB random geometric graphs is to construct a bi-partite graph on two independent Poisson point process of intensities n and cnin the unit cube. We provided almost sure asymptotic bounds for the connectivity threshold for all $c > 0$ and a suitable choice of radius cut-off functions $r_n(c)$. Further for $c < c_0$, we derived a weak law result for the largest nearest neighbor radius. This work appeared in [27].

6.4.2. Random Packing Models

Random packing models (RPM) are point processes (p.p.s) where points which "contend" with each other cannot be simultaneously present. These p.p.s play an important role in many studies in physics, chemistry, material science, forestry and geology. For example, in microscopic physics, chemistry and material science, RPMs can be used to describe systems with hard-core interactions. Applications of this type range from reactions on polymer chains, chemisorption on a single-crystal surface, to absorption in colloidial systems. In these models, each point (molecule, particle,$\cdots$) in the system occupies some space, and two points with overlapping occupied space contend with each other. Another example is the study of seismic and forestry data patterns, where RPMs are used as a reference model for the data set under consideration. In wireless communications, RPMs can be used to model the users simultaneously accessing the medium in

a wireless network using Carrier Sensing Medium Access (CSMA). In this context, each point (node, user, transmitter,$\cdots$) does not occupy space but instead generates interference to other points in the network. Two points contend with each other if either of them generates too much interference to the other. Motivated by this kind of application, we studied in [66] the generating functionals of several models of random packing processes: the classical Matérn hard-core model; its extensions, the k-Matérn models and the ∞-Matérn model, which is an example of random sequential packing process. The main new results are: 1) A sufficient condition for the ∞-Matérn model to be well-defined (unlike the other two, the ∞-Matérn model may not be well-defined on unbounded space); 2) the generating functional of the resulting point process which is given for each of the three models as the solution of a differential equation; 3) series representation and bounds on the generating functional of the packing models; 4) moment measures and other useful properties of the considered packing models which are derived from their generating functionals.

6.4.3. Extremal and Additive Matérn Point Processes

In the simplest Matérn point processes, one retains certain points of a Poisson point process in such a way that no pairs of points are at distance less than a threshold. This condition can be reinterpreted as a threshold condition on an extremal shot–noise field associated with the Poisson point process. In a joint work with P. Bermolen [Universidad de la República, Montevideo, Uruguay] [60], we studied extensions of Matérn point processes where one retains points that satisfy a threshold condition based on an *additive* shot–noise field of the Poisson point process. We provide an analytical characterization of the intensity of this class of point processes and we compare the packing obtained by the extremal and additive schemes and certain combinations thereof.

6.4.4. Spatial Birth and Death Point Processes

In collaboration with F. Mathieu [Inria GANG] and Ilkka Norros [VTT, Finland], we continued studying a new spatial birth and death point process model where the death rate is a shot noise of the point configuration. We showed that the spatial point process describing the steady state exhibits repulsion. We studied two asymptotic regimes: the fluid regime and the hard–core regime. We derived closed form expressions for the mean (and in some cases the law) of the latency of points as well as for the spatial density of points in the steady state of each regime. A paper on the matter will be presented at Infocom 13.

6.4.5. A population model based on a Poisson line tessellation

In [44], we introduce a new population model. Taking the geometry of cities into account by adding roads, we build a Cox process driven by a Poisson line tessellation. We perform several shot-noise computations according to various generalizations of our original process. This allows us to derive analytical formulas for the uplink coverage probability in each case.

6.4.6. Information Theory and Stochastic Geometry

In a joint work with V. Anantharam [UC Berkeley], we study the Shannon regime for the random displacement of stationary point processes. We currently investigate Multiple Access Channels.

6.4.7. Navigation on Point Processes and Graphs

The thesis of Mir Omid Mirsadeghi [6] studied optimal navigations in wireless networks in terms of first passage percolation on some space-time SINR graph. It established both “positive” and “negative” results on the associated percolation delay rate (delay per unit of Euclidean distance, also called time constant in the classical terminology of percolation). The latter determines the asymptotics of the minimum delay required by a packet to progress from a source node to a destination node when the Euclidean distance between the two tends to infinity. The main negative result states that the percolation delay rate is infinite on the random graph associated with a Poisson point process under natural assumptions on the wireless channels. The main positive result states that when adding a periodic node infrastructure of arbitrarily small intensity to the Poisson point process, the percolation delay rate is positive and finite.

A new direction of research was initiated aiming at defining a new class of measures on a point process which are invariant under the action of a navigation on this point process. This class of measures has properties similar to Palm measures of stationary point processes; but they cannot be defined in the classical framework of Palm measures.

6.5. Random Graphs and Combinatorial Optimization

Participants: Emilie Coupechoux, Kumar Gaurav, Mathieu Leconte, Marc Lelarge.

random graphs, combinatorial optimization, local weak convergence, diffusion, network games.

6.5.1. Matchings in infinite graphs

In [13] with Charles Bordenave [CNRS-Université de Toulouse] and Justin Salez [Université Paris 7], we proved that for any sequence of (deterministic or random) graphs converging locally, the corresponding sequence of normalized matching numbers converges, and this limit depends only on the limit of the graph sequence. In the particular case where this limit is a unimodular Galton Watson tree, we were able to compute explicitly the value for the limit of the sequence of (normalized) matching numbers. This leads to an explicit formula that considerably extends the well-known one by Karp and Sipser for Erdős-Rényi random graphs.

We considered a natural family of Gibbs distributions over matchings on a finite graph, parameterized by a single positive number called the temperature. The correlation decay technique can be applied for the analysis of matchings at positive temperature and allowed us to establish the weak convergence of the Gibbs marginal as the underlying graph converges locally. However for the zero temperature problem (i.e. maximum matchings), we showed that there is no correlation decay even in very simple cases. By using a complex temperature and a half-plane property due to Heilmann and Lieb, we were able to let the temperature tend to zero and obtained a limit theorem for the asymptotic size of a maximum matching in the graph sequence.

6.5.2. Convergence of Multivariate Belief Propagation, with Applications to Cuckoo Hashing and Load Balancing

In [58], with Laurent Massoulié [Inria-MSR], we extend the results obtained previously on the asymptotic size of maximum matchings in random graphs converging locally to Galton-Watson trees to so-called capacitated b-matchings (with non-unitary capacity at vertices as well as constraints on individual edges). Compared to the matching case, this involves studying the convergence of a message passing algorithms which transmits vectors instead of single real numbers. We also look further into an application of these results to large multiple-choice hashtables. In particular, cuckoo hashing is a popular and simple way to build a hashtable where each item is only allowed to be assigned keys within a predetermined, random subset of all keys. In this context, it is important to determine the load threshold under which cuckoo hashing will succeed with high probability in building such a hashtable. The results on the density of maximum capacitated b-matchings allow to determine this threshold.

6.5.3. A new approach to the orientation of random hypergraphs

A h-uniform hypergraph $H = (V, E)$ is called (l, k)-orientable if there exists an assignment of each hyperedge e to exactly l of its vertices such that no vertex is assigned more than k hyperedges. Let $H_{n,m,h}$ be a hypergraph, drawn uniformly at random from the set of all h-uniform hypergraphs with n vertices and m edges. In [41], we determine the threshold of the existence of a (l, k)-orientation of $H_{n,m,h}$ for $k \geq 1$ and $h > l \geq 1$, extending recent results motivated by applications such as cuckoo hashing or load balancing with guaranteed maximum load. Our proof combines the local weak convergence of sparse graphs and a careful analysis of a Gibbs measure on spanning subgraphs with degree constraints. It allows us to deal with a much broader class than the uniform hypergraphs.

6.5.4. Bipartite graph structures for efficient balancing of heterogeneous loads

In [40], with Laurent Massoulié [Inria-MSR], we look into another application of the results on the asymptotic maximum size of b-matchings to large scale distributed content service platforms, such as peer-to-peer video-on-demand systems. In this context, the density of maximum b-matchings corresponds to the maximum fraction of simultaneously satisfiable requests, when the service resources are limited and each server can only handle requests for a predetermined subset of the contents which it has stored in memory. An important design aspect of such systems is the content placement strategy onto the servers depending on the estimated content popularities; the results obtained allow to characterize the efficiency of such placement strategies and the optimal strategies in the limit of large storage capacity at servers are determined.

6.5.5. Flooding in Weighted Random Graphs

In a joint work [8] with Hamed Amini [EPFL] and Moez Draief [Imperial College London], we studied the impact of the edge weights on distances in diluted random graphs. We interpret these weights as delays, and take them as i.i.d exponential random variables. We analyzed the edge flooding time defined as the minimum time needed to reach all nodes from one uniformly chosen node, and the edge diameter corresponding to the worst case edge flooding time. Under some regularity conditions on the degree sequence of the random graph, we showed that these quantities grow as the logarithm of n, when the size of the graph n tends to infinity. We also derived the exact value for the prefactors.

These allowed us to analyze an asynchronous randomized broadcast algorithm for random regular graphs. Our results show that the asynchronous version of the algorithm performs better than its synchronized version: in the large size limit of the graph, it will reach the whole network faster even if the local dynamics are similar on average.

6.5.6. Upper deviations for split times of branching processes

In [9], upper deviation results are obtained for the split time of a supercritical continuous-time Markov branching process. More precisely, with Hamed Amini [EPFL], we establish the existence of logarithmic limits for the likelihood that the split times of the process are greater than an identified value and determine an expression for the limiting quantity. We also give an estimation for the lower deviation probability of the split times which shows that the scaling is completely different from the upper deviations.

6.5.7. Epidemics in random clustered networks

In [54], 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 this paper, 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.

6.5.8. Leveraging Side Observations in Stochastic Bandits

The paper [37] considers stochastic bandits with side observations, a model that accounts for both the exploration/exploitation dilemma and relationships between arms. In this setting, after pulling an arm i, the decision maker also observes the rewards for some other actions related to i. We will see that this model is

suited to content recommendation in social networks, where users' reactions may be endorsed or not by their friends. We provide efficient algorithms based on upper confidence bounds (UCBs) to leverage this additional information and derive new bounds improving on standard regret guarantees. We also evaluate these policies in the context of movie recommendation in social networks: experiments on real datasets show substantial learning rate speedups ranging from 2.2x to 14x on dense networks.

6.5.9. Universality in Polytope Phase Transitions and Message Passing Algorithms

In [28], with Mohsen Bayati and Andrea Montanari [Stanford], we consider a class of nonlinear mappings F in R^N indexed by symmetric random matrices A in $R^{N\times N}$ with independent entries. Within spin glass theory, special cases of these mappings correspond to iterating the TAP equations and were studied by Erwin Bolthausen. Within information theory, they are known as 'approximate message passing' algorithms. We study the high-dimensional (large N) behavior of the iterates of F for polynomial functions F, and prove that it is universal, i.e. it depends only on the first two moments of the entries of A, under a subgaussian tail condition. As an application, we prove the universality of a certain phase transition arising in polytope geometry and compressed sensing. This solves -for a broad class of random projections- a conjecture by David Donoho and Jared Tanner.

6.5.10. Far-out Vertices In Weighted Repeated Configuration Model

In [34] we consider an edge-weighted uniform random graph with a given degree sequence (Repeated Configuration Model) which is a useful approximation for many real-world networks. It has been observed that the vertices which are separated from the rest of the graph by a distance exceeding certain threshold play an important role in determining some global properties of the graph like diameter, flooding time etc., in spite of being statistically rare. We give a convergence result for the distribution of the number of such far-out vertices. We also make a conjecture about how this relates to the longest edge of the minimal spanning tree on the graph under consideration.

ALPAGE Project-Team

6. New Results

6.1. Advances in symbolic and hybrid parsing with DyALog and FRMG

Participants: Éric Villemonte de La Clergerie, François Barthélemy, Julien Martin.

Within the team is developed a wide-coverage French meta-grammar (FRMG) and a efficient hybrid TAG/TIG parser based on the DyALog logic programming environment [120] and on the Le*fff* morphological and syntactic lexicon [105]. It relies on the notion of factorized grammar, themselves generated from a representation that lies at a higher level of abstraction, named Meta-Grammars [122]. At that level, linguistic generalizations can be expressed, which in turn makes it possible to transfer meta-grammars from one language to a closely related one. The hybrid TAG/TIG parser generator itself implements all kinds of parsing optimizations: lexicalization (in particular via hypertags), left-corner guiding, top/bottom feature analysis, TIG analysis (with multiple adjoining), and others. The recent evolutions go towards an hybridization with statistical approaches.

6.1.1. Tuning FRMG's disambiguation mechanism

Continuing works initiated in 2011 on the exploitation of the dependency version of the French TreeBank (FTB), Éric de La Clergerie has explored the tuning of FRMG's rule base disambiguation mechanism using a larger set of features and weight learned from the FTB. In 2011, this approach led to on improvement from 82.31% to 84.54% in terms of accuracy (LAS - Labelled Attachment Score) on the test part of the FTB. By increasing the set of features, in particularly using higher-order dependency features (on parent edge and sibling edges), and a better understanding of the iterative tuning mechanism, it was possible to reach 85.95% LAS. This tuning mechanism is based on the idea of adding or subtracting some weight to a disambiguation rule given some specific contexts (provided by the features), where the delta is progressively learned from the accuracy of the disambiguation rule in terms of edge selection or rejection. The learning algorithm presents some relationships with the perceptron approach, but the use of a more standard implementation of the perceptron led to less interesting gains.

During the same time, the coverage of FRMG was improved (to reach for instance 94% of full parses on the FTB).

6.1.2. Synchronous Tree-Adjoining Grammars

A preliminary work has been done to implement *Synchronous Tree-Adjoining Grammars* (STAGs) in DyALog, relying on the notion of *Thread Automata* [119]. Synchronous Tree Adjoining Grammars is an instance of formalism where the order of the components of a tree structure is not fully determined. This leads to combinatorial alternatives when parsing, while a tree-structure corresponding to the input string has to be build. A specific front-end has been written to implement STAGs. The work on the back-end is still in progress, with the goal to have a common intermediate representation for several mildly context-sensitive formalisms where some node operations non-deterministically pick a node out of a finite set of nodes. STAGs are an instance of such formalisms, Multi-Component Tree Adjoining Grammars (MCTAGs) are another instance. The intermediate representation consists in Thread Automata (TA), an extension of Push-Down Automata where several threads of computations are considered and only one is active at any time.

6.1.3. Adding weights and probabilities to DyALog

Weights can already be used during the disambiguation phase of the FRMG parser, implemented in DyALog. However, a deeper implementation of weights and probabilities in DyALog was initiated in 2012 by Julien Martin during his Master internship. By enriching the structure of the backpointers (relating the items to their parent items), it is now possible to maintain an ordered weighted list of derivations, to update the scheduling of items wrt their weight, to update the weights of all the descendants of an item I when updating I's weight. The motivation is of course to be able to favor the best analysis first during parsing. A second objective (which has

been implemented) is the possibility to extract the n-best parses after parsing (but keeping a shared derivation forest). A third objective, remaining to be done, is related to the use of beam search techniques to prune the search space during parsing. A longer-term objective is the abstraction of this work to be able to work on semi-rings.

6.2. Tree transformation

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

In 2011, the conversion of native FRMG dependencies into the CONLL dependency scheme was the occasion to explore new ideas about tree transformation (for dependencies), based on the notion of two-level transformation with a first level relying on local transformation rules and a second level being controlled by constraints carried by the first level edges. During his Master internship, Corentin Ribeyre has formalized and re-implemented this approach in a more systematic and generic way. This work was also completed by the use of example-based learning techniques to quickly learn the local transformation rules of the first level. The line of research is motivated by possibility to quickly develop a reduced set of transformation rules (thanks to the examples and the constraint level) for a large variety of applications, such as information extration but also conversion toward a deep syntax level or a shallow semantic level. A poster paper was presented at TAG+11 [29].

6.3. lexical knowledge acquisition and visualization

Participants: Éric Villemonte de La Clergerie, Mickael Morardo, Benoît Sagot.

In relation with our collaboration with Lingua & Machina (cf section 4.4), Mikael Morardo has enriched the interfaces of the WEB platform Libellex for the visualization and validation of more complex lexical resource. In particular, the focus has been on the development of a graph-based view with the javascript Library `d3.js` to represent large lexical networks. The current implementation is powerful enough to deal with large networks of several teens of thousands of connections, allowing the visualization of fragments of the network and an easy navigation. Because the graph-view proved to be both intuitive and efficient, the previous list-based view for terminology was partially re-implemented in the new graph-view. It was also extended for visualizing and validating more complex lexical networks, like the French Wordnet WOLF coupled with the original English WordNet (cf 5.9).

The graph-based view was used to explore several networks built using Harris' distributional hypothesis (through a clustering algorithm) on the output of FRMG for several corpora. Because terminology was now be visualized at the same time, the clustering algorithm was modified to be able to take into account a list of terms (also automatically extracted from the parsed corpora) .

6.4. Advances in statistical parsing

Participants: Marie Candito, Benoît Crabbé, Djamé Seddah, Enrique Henestroza Anguiano.

6.4.1. Statistical Parsing

We have achieved **state-of-the art results for French statistical parsing**, adapting existing techniques for French, a language with a morphology richer than English, either for constituency parsing [110], [113] or dependency parsing [68]. We made available The Bonsai parsing chain [1] (cf. 5.4), that gathers preprocessing tools and models for French dependency parsing into an easy-to-use parsing tool for French. We designed our parsing pipeline with modularity in mind: our parsing models are interchangeable. For instance, dependencies output can either be generated from a PCFG-LA based parser associated with a functional role labeler or from any dependency parsers trained on our dependency treebank [68]. Tokens can either be raw words, POS tagged lemmas or word clusters [69].

[1] http://alpage.inria.fr/statgram/frdep/fr_stat_dep_parsing.html

We have innovated in the tuning of tagsets to optimize both grammar induction and unknown word handling [75], thus providing the best parsing models for French [111]. Then we have contributed on three main points:

1. conversion of the French Treebank [55] used as constituency training data into a dependency treebank [4], which is now used by several teams for dependency parsing;
2. an original method to reduce lexical data sparseness by replacing tokens by unsupervised word clusters, or morphological clusters [64], [112];
3. a postprocessing step that uses specialized statistical models for parse correction [81].

For the last 18 to 12 months, we have been increasingly focused in increasing the robustness of our parsing models by (a) validating our approach on other morphologically-rich languages; (b) other domains and (c) on user generated content. All of those challenging the current state-of-the-art in statistical parsing.

6.4.2. Multilingual parsing

Applying the techniques we developed for reducing lexical data, which is commonly found in morphologically-rich languages (MRLs) and optimizing the POS tagset, we integrated lexical information through data driven lemmatisation [112] and POS tagging [79]. This provided state-of-the-art results in parsing Romance languages such as Italian [35] and Spanish [26]. In the latter case, we mixed the outputs of two morphological analyzers and generated a version of the treebank where each morphological gold information was replaced by a predicted one. Relying on a rich lexicon developed within the Alexina framework (cf. 5.8) and accurate morphological treatment (cf. 6.5), this method brings more robustness to treebank-based parsing models.

6.4.3. Out-of-domain parsing : resources and parsing techniques

Statistical parsing is known to lead to parsers that exhibit quite degraded performance on input text that varies from the sentences used for training. Alpage has devoted a major effort on providing both evaluation resources and parser adaptation techniques, to increase robustness of statistical parsing for French. We have investigated several degrees of distance between the training corpus, the French Treebank, which is made of sentences from the *Le Monde* newspaper: we first focused on parsing well-edited texts, but from domains with varying difference with respect to the national newspaper *Le Monde* type of text. We then turned our attention to parsing user-generated content, hence potentially not only from a different domain than news, but also with great "noise" with respect to well-edited texts, and extremely divergent linguistic phenomena (see next subsection). As far as out-of-domain well-edited text, we have supervised the annotation and release of the **Sequoia Treebank** [47] (https://www.rocq.inria.fr/alpage-wiki/tiki-index.php?page=CorpusSequoia), a corpus of 3200 sentences annotated for part-of-speech and syntactic structure, from four subdomains : sentences from the regional newspaper *L'Est Républicain*, from the French Wikipedia, from the Europarl Corpus (European parliamentary debates), and from reports of the European Medicine Agency. We have proposed a word clustering technique, with clusters computed over a "*bridge*" corpus that couples indomain and target domain raw texts, to improve parsing performance on target domain, without degrading performance on indomain texts (contrary to usual adaptation techniques such as self-training). Preliminary experiments were performed on the biomedical domain only [67] and confirmed on the whole Sequoia Treebank [47].

6.4.4. Robust parsing of user-generated content

Until very recently out-of-domain text genres that have been prioritized have not been Web 2.0 sources, but rather biomedical texts, child language and general fiction (Brown corpus). Adaptation to user-generated content is a particularly difficult instance of the domain adaptation problem since Web 2.0 is not really a domain: it consists of utterances that are often ungrammatical. It even shares some similarities with spoken language [116]. The poor overall quality of texts found on such media lead to weak parsing and even POS-tagging results. This is because user-generated content exhibits both the same issues as other out-of-domain data, but also tremendous issues related to tokenization, typographic and spelling issues that go far beyond what statistical tools can learn from standard corpora. Even lexical specificities are often more challenging than on edited out-of-domain text, as neologisms built using productive morphological derivation, for example, are less frequent, contrarily to slang, abbreviations or technical jargon that are harder to analyze and interpret automatically.

In order to fully prepare a shift toward more robustness, we started to develop a richly annotated corpus of user-generated French text, the French Social Media Bank, which includes not only POS, constituency and functional information, but also a layer of "normalized" text[37]. This corpus is fully available and constitutes the first data set on Facebook data and the first instance of user generated content for an MRL.

Besides delivering a new data set, our main purpose here is to be able to compare two different approaches to user-generated content processing: either training statistical models on the original annotated text, and use them on raw new text; or developing normalization tools that help improving the consistency of the annotations, train statistical models on the normalized annotated text, and use them on normalized texts (before un-normalizing them).

However, this raises issues concerning the normalization step. A good sandbox for working on this challenging task is that of POS-tagging. For this purpose, we did leverage Alpage's work on MElt, a state-of-the art POS tagging system [15]. A first round of experiments on English have already led to promising results during the shared task on parsing user-generated content organized by Google in May 2012 [93], as Alpage was ranked second and third [38]. For achieving this result, we brought together a preliminary implementation of a normalization wrapper around the MElt POS tagger followed by a state-of-the art statistical parser improved by several domain adaptation techniques originally developed for parsing edited out-of-domain texts (cf. previous section).

One of our objectives is to generalize the use of the normalization wrapper approach to both POS tagging and parsing, for English and French, in order to improve the quality of the output parses. However, this raises several challenges: non-standard contractions and compounds lead to unexpected syntactic structures. A first round of experiments on the French Social Media Bank showed that parsing performance on such data are much lower than expected. This is why, we are actively working to improve on the baselines we established on that matter.

6.4.5. Precise recovery of unbounded dependencies

We focused on a linguistic phenomena known as long-distance dependencies. These are dependencies involved a fronted element that depends on a head that is potentially embedded in the clause the element is in front of. This embedding make such dependencies very hard to recover for a parser. Though this phenomena is rare, the corresponding dependencies are generally part of predicate-argument structures, and are thus very important to recover for downstream semantic applications. We have assessed the low parsing performance of long-distance dependencies (LDDs) for French, proposed an explicit annotation of such dependencies in the French Treebank and the Sequoia Treebank, and evaluated several parsing architectures with the aim of maintaining high general performance and good performance on LDDs [22]. We found that using a non-projective parser helps for LDDs but degrades overall performance, while using pseudo-projective parsing [88] which transforms in a reversible way a non-projective treebank into a projective one) is the best strategy, in order to take advantage of the better performance of projective parsers.

6.5. Computational morphology and automatic morphological analysis

Participants: Benoît Sagot [correspondant], Marion Baranes, Virginie Mouilleron, Damien Nouvel.

Since 2011 and, Alpage members have started interacting with formal morphologists for taking part in the development and implementation of new morphological models and resources. Concerning inflectional morphology, this work has led to new versions of the morphological layer of the ALEXINA formalism, to new ALEXINA lexicons for several languages of choice (Kurdish languages and German, as mentioned above, but also Maltese and Latin, see the section on ALEXINA), and to studies about the quantitative assessment of morphological complexity, currently an active area of research in morphology, have been pursued following previous work published in 2011 [109], [126]. Concerning constructional morphology (derivation, composition) and borrowings, studies and experiments have been carried out in the context of the ANR EDyLex project and that of the collaboration with *viavoo* [45], following here as well experiments carried out in 2011 [124], [115], [127].

6.6. Advances in lexical morphology and syntax

Participants: Benoît Sagot [correspondant], Laurence Danlos, Éric Villemonte de La Clergerie.

The Alexina framework (cf. 5.8) [105] has been developed and used for developing various lexicons, in particular the Le*fff*, that are used in many tools such as POS-taggers [15] and parsers.

In 2012, the new developments within Alexina have been fourfold:

- A large amount of work has been made for developing a new morphological layer to Alexina, in collaboration with a specialist of formal morphology.
- In the context of this collaboration, new Alexina lexicons have been developed with a special focus on linguistic relevance and exhaustivity within a well-defined subset of lexical entries (e.g., Latin verbs, 1st-binyan Maltese verbs).
- The development of a new large-scale NLP-oriented Alexina lexicon has been initiated, namely that of DeLex, an Alexina lexicon for German. It is currently restricted to the morphological layer (no valency information yet) but already generates 2 million inflected lexical entries. The underlying morphological grammar makes use of the new morphological layer mentioned above.
- Following previous work, merging experiments between syntactic resources and the Le*fff* [30] and comparison experiments between such resources and the Le*fff* as reference lexicon for the FRMG parser have been carried out [43]. In the latter series of experiments, the Le*fff* has proven better, or rather more suitable, that other (converted) resources.

6.7. Named Entity Recognition and Entity Linking

Participants: Rosa Stern, Benoît Sagot.

Identifying named entities is a widely studied issue in Natural Language Processing, because named entities are crucial targets in information extraction or retrieval tasks, but also for preparing further NLP tasks (e.g., parsing). Therefore a vast amount of work has been published that is dedicated to named entity *recognition*, i.e., the task of identification of named entity *mentions* (spans of text denoting a named entity), and sometimes *types*. However, real-life applications need not only identify named entity mentions, but also know which real entity they refer to; this issue is addressed in tasks such as knowledge base population with entity resolution and linking, which require an inventory of entities is required prior to those tasks in order to constitute a reference.

6.7.1. Cooperation of symbolic and statistical methods for named entity recognition and typing

Named entity recognition and typing is achieved both by symbolic and probabilistic systems. We have performed an experiment [62] for making the rule-based system NP, SXPipe's high-precision named entity recognition system developed at Alpage on AFP news corpora and which relies on the *Aleda* named entity database, interact with LIANE, a high-recall probabilistic system developed by Frédéric Béchet (LIF) and trained on oral transcriptions from the ESTER corpus. We have shown that a probabilistic system such as LIANE can be adapted to a new type of corpus in a non-supervised way thanks to large-scale corpora automatically annotated by NP. This adaptation does not require any additional manual annotation and illustrates the complementarity between numeric and symbolic techniques for tackling linguistic tasks.

6.7.2. Nomos, a statistical entity linking system

For information extraction from news wires, entities such as persons, locations or organizations are especially relevant in a knowledge acquisition context. Through a process of named entity recognition and entity linking applied jointly, we aim at the extraction and complete identification of these relevant entities, which are meant to enrich textual content in the form of *metadata*. In order to store and access extracted knowledge in a structured and coherent way, we aim at populating an ontological reference base with these metadata. We have pursued our efforts in this direction, using an approach where NLP tools have early access to Linked Data resources and thus have the ability to produce metadata integrated in the Linked Data framework. In particular, we have studied how the entity linking process in this task must deal with noisy data, as opposed to the general case where only correct entity identification is provided.

We use the symbolic named entity recognition system NP, a component of SXPipe, and use it as a mention detection module. Its output is then processed through our entity linking system, which is based on a supervised model learned from examples of linked entities. Since our named entity recognition is not deterministic, as opposed to other entity linking tasks where the gold named entity recognition results are provided, it is configured to remain ambiguous and non-deterministic, i.e., its output preserves a number of ambiguities which are usually resolved at this level. In particular, no disambiguation is made in the cases of multiple possible mentions boundaries (e.g., *{Paris}+{Hilton}* vs. *{Paris Hilton}*). In order to cope with possible false mention matches, which should be discarded as linking queries, the named entity recognition output is made more ambiguous by adding a *not-an-entity* alternative to each mention's candidate set for linking. The entity linking module's input therefore consists in multiple possible readings of sentences. For each reading, this module must perform entity linking on every possible entity mention by selecting their most probable matching entity. Competing readings are then ranked according to the score of entities (or sequence of entities) ranked first in each of them. The reading with no entity should also receive a score in order to be included in the ranking. The motivation for this joint task lies in the frequent necessity of accessing contextual and referential information in order to complete an accurate named entity recognition; thus the part where named entity recognition usually resolves a number of ambiguities is left for the entity linking module, which uses contextual and referential information about entities.

We have realized a first implementation of our system, as well as experiments and evaluation results. In particular, when using knowledge about entities to perform entity linking, we discuss the usefulness of domain specific knowledge and the problem of domain adaptation.

In 2012, improvements have been made to Nomos by combining the NP named entity detection module with LIANE, a probabilistic system developed by Frédéric Béchet (LIF) in order to better predict possible false matches. The linking step has also been enriched with the use of a more complete and autonomous knowledge base derived from Wikipedia, as well as new parameters and ranking functions for the prediction of the mention/entity alignment.

In the context of this linking task for the processing of AFP corpora and content enrichment with metadata, we conducted a deep study of Semantic Web recent developments and especially of the Linked Data initiatives in order to consider the integration of AFP metadata in these knowledge representation frameworks. On this topic as well as the enlarged view of entity linking for semantic annotation of textual content, discussions have taken place with Eric Charton (CRIM, Montréal, Canada) during 2012 Fall.

The Nomos system as well as the general process of content enrichment with metadata and reference base population has been presented at a dedicated workshop at NAACL in June 2012 (AKBC-WEKEX 2012).

6.8. Advances in lexical semantics

Participants: Benoît Sagot [correspondant], Marion Richard, Sarah Beniamine.

In 2012, several contributions to the WOLF have been finalized and/or published. In particular, various successful attempts to enhance the coverage of the WOLF have been integrated within the master resource [23], [19], [31], [24]. A more original work has also been achieved, targeted at improving the precision of the resource by automatically detecting probable outliers [32]. This latter work has been integrated within the dedicated sloWTool platform, and these outliers partly validated by Slovene students of Romance studies. In parallel, a medium-scale manual validation effort has been achieved at Alpage thanks to the work of two Master students funded by the ANR EDyLex project, which has led to the validation of a vast majority of so-called "basic" synsets, i.e., what can be expected to be the most useful part of the resource.

The result of all this work has been integrated in a preliminary first non-alpha version of the WOLF, version WOLF 1.0b.

6.9. Techniques for transferring lexical resources from one language to a closely-related one

Participants: Yves Scherrer, Benoît Sagot.

Developing lexical resources is a costly activity, which means that large resources only exist for a small number of languages. In our work, we address this issue by transferring linguistic annotations from a language with large resources to a closely related language which lacks such resources. This research activity, funded by the Labex EFL, has started in October 2012.

First results include the development of a method to create bilingual dictionaries without any parallel data, depending solely on surface form similarities and their regularities. The resulting bilingual dictionaries are used to transfer part-of-speech annotations from one language to the other. At the moment, our methods are being tested with Wikipedia texts from various languages and dialects closely related to German, such as Dutch and Pfälzisch. We plan to extend this work to data from other language groups and to other types of linguistic annotations, for instance syntactic or semantic resources.

6.10. Modelling the acquisition of linguistic categories by children

Participants: Benoît Crabbé, Luc Boruta, Isabelle Dautriche.

This task breaks in two sub-tasks: acquisition of phonemic categories, and acquisition of syntactic categories.

Although we are only able to distinguish between a finite, small number of sound categories – i.e., a given language's phonemes – no two sounds are actually identical in the messages we receive. Given the pervasiveness of sound-altering processes across languages – and the fact that every language relies on its own set of phonemes – the question of the acquisition of allophonic rules by infants has received a considerable amount of attention in recent decades. How, for example, do English-learning infants discover that the word forms [kæt] and [kat] refer to the same animal species (i.e. *cat*), whereas [kæt] and [bæt] (i.e. *cat*∼*bat*) do not? What kind of cues may they rely on to learn that [sɪŋkɪŋ] and [θɪŋkɪŋ] (*sinking*∼*thinking*) can not refer to the same action? The work presented in this dissertation builds upon the line of computational studies initiated by [90], wherein research efforts have been concentrated on the definition of sound-to-sound dissimilarity measures indicating which sounds are realizations of the same phoneme. We show that solving Peperkamp et al.'s task does not yield a full answer to the problem of the discovery of phonemes, as formal and empirical limitations arise from its pairwise formulation. We proceed to circumvent these limitations, reducing the task of the acquisition of phonemes to a partitioning-clustering problem and using multidimensional scaling to allow for the use of individual phones as the elementary objects. The results of various classification and clustering experiments consistently indicate that effective indicators of allophony are not necessarily effective indicators of phonemehood. Altogether, the computational results we discuss suggest that allophony and phonemehood can only be discovered from acoustic, temporal, distributional, or lexical indicators when—on average—phonemes do not have many allophones in a quantified representation of the input. This subtask has seen the Phd defense of Luc Boruta whose Phd thesis : *"Indicators of allophony and phonemehood"* was successfully defended in September 2012.

As for syntactic categorization, the task is concerned with modelling and implementing psychologically motivated models of language treatment and acquisition. Contrary to classical Natural Language Processing applications, the main aim was not to create engineering solutions to language related tasks, but rather to test and develop psycholinguistic theories. In this context, the study was concerned with the question of learning word categories, such as the categories of Noun and Verb. It is established experimentally that 2-year-old children can identify novel nouns and verbs. It has been suggested that this can be done using distributional cues as well as prosodic cues. While the plain distributional hypothesis had been tested quite extensively, the importance of prosodic cues has not been addressed in a computational simulation. We provided a formulation for modelling this hypothesis using unsupervised and semi-supervised forms of Bayesian learning (EM) both offline and online. This activity started with the master thesis of A. Gutman and has seen this year the start of a new Phd student : I. Dautriche.

6.11. Modelling and extracting discourse structures

Participants: Laurence Danlos, Charlotte Roze.

6.11.1. Lexical semantics of discourse connectives

Discourse connectives are words or phrases that indicate senses holding between two spans of text. The theoretical approaches accounting for these senses, such as text coherence, cohesion, or rhetorical structure theory, share at least one common feature: they acknowledge that many connectives can indicate different senses depending on their context. LEXCONN is a lexical database for French connectives [16].

The French connectives "*en réalité*" and "*en effet*" have been the topic of numerous studies but none of them was formalized. [53] gives a formalization of the conditions the two arguments of these connectives should meet. This formalization is based on factivity information as modeled in the FactBank corpus developed by Roser Sauri.

Sometimes, the sense of connectives is unique but its arguments are hard to determine. In particular, the second argument of an adverbial connective is not always equivalent to its syntactic arguments. This raises problems at the syntax-semantics interface which are described in [52]. The method to handle theses problems in a discursive parser will be studied in the ANR project POLYMNIE, which is headed by Sylvain Podogolla (Inria Lorraine) and which started in October 2012.

6.11.2. Discursive annotation

We plan to annotate the French corpus FTB (French Tree Bank) at the discursive level, in order to obtain the FDTB (French Discourse Tree Bank). The methodology that will be used is close to the one used in the PDTB (Penn Discourse Tree Bank). The first steps of this long term project are presented in [48], [49], [51].

This work is based on a new hierarchy of discourse relations and this new hierarchy was presented at an European workshop organized by the project MULDICO.

6.12. Modelling word order preferences in French

Participants: Juliette Thuilier, Benoît Crabbé, Margaret Grant.

We study the problem of choice in the ordering of French words using statistical models along the lines of [60] and [61]. This work aims at describing and model preferences in syntax, bringing additional elements to Bresnan's thesis, according to which the syntactic competence of human beings can be largely simulated by probabilistic models. We previously investigated the relative position of attributive adjectives with respect to the noun.

This year has seen the Phd thesis defense of Juliette Thuilier in September 2012.

In collaboration with Anne Abeillé (Laboratoire de Linguistique Formelle, Université Paris 7), we extended our corpora study with psycholinguistic questionnaires, in order to show that statistical models are reflecting some linguistic knowledge of French speakers. The preliminary results confirm that animacy is not a relevant factor in ordering French complements.

As regards to corpus work, we are extending the database with spontaneous speech corpora (CORAL-ROM and CORPAIX) and a wider variety of verbal lemmas, in order to enhance sample representativeness and statistical modelling. This activity has lead to the development of an extension of the French Treebank for oral corpora (approx 2000 sentences).

In a cross-linguistic perspective, we plan to strengthen the comparison with the constraints observed in other languages such as English or German with the recruitment of a new postdoc arriving at the beginning of 2013.

As can be seen from the outline above, this line of research brings us closer to cognitive sciences. We hope, in the very long run, that these investigations will bring new insights on the design of probabilistic parsers or generators. In NLP, the closest framework implementing construction grammars is Data Oriented Parsing (DOP).

AXIS Project-Team

5. New Results

5.1. Introduction

As planned, our new results are splitted into our three sub-objectives as introduced below:

5.1.1. Mining for Knowledge Discovery in Information Systems

This year we get six main results: one related to how to integrate domain knowledge in a multi-view KDD process (cf. section 5.2.4), two on new KDD methods involving clustering (cf. sections 5.2.3 and 5.2.2), one on the construction of hierarchical structures of concepts in the field of e-tourism (cf. section 5.2.6), one on partitioning objects taking into account simultaneously their relational descriptions given by multiple dissimilarity matrices (cf. section 5.2.1), and finally improvment of our work on critical edition of Sanskrit texts (cf. section 5.2.5).

- Zhang based on his thesis (2010) has published this year his work on modeling and clustering users with evolving profiles in usage streams [32]. This paper will propose three models to summarize bi-streaming data, which are the batch model, the Evolving Objects (EO) model and the Dynamic Data Stream (DDS) model. Through creating, updating and deleting user profiles, the models summarize the behaviours of each user as an object. Based on these models, clustering algorithms are employed to identify the user groups. The proposed models are tested on a real-world data set showing that the DDS model can summarize the bi-streaming data efficiently and effectively, providing better basis for clustering user profiles than the other two models.
- The work described in 2011(see our AxIS annual report) on critical edition of Sanskrit texts and submitted as a paper at the Cicling 2012 conference has been accepted [21].
- A past work accepted in an international journal with A. Ciampi and colleagues [16].
- One article in an international conference on functional data analysis issued from a collaboration with F. Rossi [40].
- Two articles have been deposit in the Computing Research Repository (CoRR): one on clustering Dynamic Web Usage data [65] from A. Da Silva's thesis and one on functional data analysis [66].

5.1.2. Information and Social Networks Mining for Supporting Information Retrieval

This year, we pursued two main works on clustering methods:

- the detection of communities in a social network (graph extracted from relationnal data) (cf. section 5.3.1),
- the improvment of our dynamic hard clustering method for relational data (cf. section 5.3.2).

5.1.3. Multidisciplinary Research For Supporting User Oriented Innovation

With the expansion of the innovation community beyond the firm's boundaries (the so-called "open innovation") a lot of changes have been introduced in design and evaluation processes : the users can become co-designers, HCI design and evaluation focus is no longer placed on usability only but also on the whole user experience, experimentations take place out of lab with large number of heterogeneous people instead of carefully controlled panels of users ... All these deep changes require improvements of existing practices, methods and tools for the design / evaluation of information systems as well as for usage analysis. This evolution calls also for a structured user centered methodology (methods and ICT tools) to deal with open innovation. Various different disciplines and trends are dedicated in understanding user behavior on Internet and with Digital Technologies, notably Human Computer Interaction community (HCI), CSCW, Workplace Studies, Distributed Cognition and Data Mining. Our contribution to open innovation research keeps its focus on usage analysis for design, evaluation and maintenance of information systems and our activities this year, as indicated in our roadmap presented at the Inria theme evaluation (2011) have been conducted both breadth wise and in depth with two main objectives :

- Improving, designing and evaluation support tools for innovation,
- Development of the FocusLab platform.

The research was conducted along three focus:

- Extension of usability methods and models (cf. section 5.4),
- Designing and evaluating user experience in the context of a living lab process (cf. section 5.5),
- FocusLab Platform (cf. section 5.6).

Let us note one research work related to Living labs done in 2011 and published in 2012 [26].

5.2. Mining for Knowledge Discovery in Information Systems

5.2.1. Clustering on Multiple Dissimilarity Matrices

Participants: Yves Lechevallier, F.A.T. de Carvalho, Guillaume Pilot, Brigitte Trousse.

In [17], we introduce hard clustering algorithms that are able to partitioning objects taking into account simultaneously their relational descriptions (relations + values) given by multiple dissimilarity matrices. The aim is to obtain a collaborative role of the different dissimilarity matrices in order to obtain a final consensus partition. These matrices could have been generated using different sets of variables and a fixed dissimilarity function or using a fixed set of variables and different dissimilarity functions, or using different sets of variables and dissimilarity functions.

During 2012 we show interest and disadvantages of these approaches to classifying curves with a Urso and Vichi distance based on the mathematical properties of curves (first derivative and second). The curves are issued from temperature sensors placed in 40 offices during one year (See section 6.1.3). This period was divided into the periods before and after challenge and the challenge period. During the challenge period the occupants had information by bonus / malus messages on energy consumption [34].

5.2.2. Web Page Clustering based on a Community Detection Algorithm

Participants: Yves Lechevallier, Yacine Slimani.

Extracting knowledge from Web user's access data in Web Usage Mining (WUM) process is a challenging task that is continuing to gain importance as the size of the Web and its user-base increase. That is why meaningful methods have been proposed in the literature in order to understand the behaviour of the user in the Web and improve the access modes to information. We pursued our previous work [102] and defined a new approach of knowledge extraction using graph theory. which is described in [29].

This work is done in collaboration with the laboratory LRIA At the Ferhat Abbas University, Sétif, Algérie.

5.2.3. Multi-criteria Clustering with Weighted Tchebycheff Distances for Relational Data

Participants: F.A.T. de Carvalho, Yves Lechevallier.

The method described in [27] uses a nonlinear aggregation criterion, weighted Tchebycheff distances, more appropriate than linear combinations (such as weighted averages) for the construction of compromise solutions. We obtain a partition of the set of objects, the prototype of each cluster and a weight vector that indicates the relevance of each criterion in each cluster. Since this is a clustering algorithm for relational data, it is compatible with any distance function used to measure the dissimilarity between objects.

5.2.4. Knowledge management in Multi-View KDD Process

Participant: Brigitte Trousse.

E.L. Moukhtar Zemmouri, in the context of his PhD thesis supervised by Hicham Behja, A. Marzark and B. Trousse pursued his work based on a Viewpoint Model in the context of a KDD process [30], [19].

Knowledge Discovery in Databases (KDD) is a highly complex, iterative and interactive process aimed at the extraction of previously unknown, potentially useful, and ultimately understandable patterns from data. In practice, a KDD process (data mining project according to CRISP-DM vocabulary) involves several actors (domain experts, data analysts, KDD experts, etc.) each with a particular viewpoint. We define a multi-view analysis as a KDD process held by several experts who analyze the same data with different viewpoints.
We propose to support users of multi-view analysis through the development of a set of semantic models to manage knowledge involved during such an analysis. Our objective is to enhance both the reusability of the process and coordination between users.
To do so, we propose first a formalization of viewpoint in KDD and a Knowledge Model that is "a specification of the information and knowledge structures and functions involved during a multi-view analysis". Our formalization, using OWL ontologies, of viewpoint notion is based on CRISP-DM standard through the identification of a set of generic criteria that characterize a viewpoint in KDD. Once instantiated, these criteria define an analyst viewpoint. This viewpoint will guide the execution of the KDD process, and then keep trace of reasoning and major decisions made by the analyst.
Then, to formalize interaction and interdependence between various analyses according to different viewpoints, we propose a set of semantic relations between viewpoints based on goal-driven analysis. We have defined equivalence, inclusion, conflict, and requirement relations. These relations allow us to enhance coordination, knowledge sharing and mutual understanding between different actors of a multi-view analysis, and reusability in terms of viewpoint of successful data mining experiences within an organization.

5.2.5. Critical Edition of Sanskrit Texts

Participants: Yves Lechevallier [correspondant], Marc Csernel, Ehab Assan.

With the help of Ehab Assan we improved the prototype made last year by Nicolas Bèchet (cf. 2011 AxIS activity report,[21]). It is now included in the construction process of critical editions of Sanskrit texts. Ehab also added LaTeX output to the process, we now have paper as well as Web output. It was possible to present these new features [33], [36] at the 13th International Conference on Intelligent Text Processing and Computational Linguistics (CICLing) in Delhi.

5.2.6. Construction and Settlement of hierarchical Structures of Concepts in E-tourism

Participant: Yves Lechevallier.

The work of Nicolas Bechet (AxIS member in 2011) and Yves Lechevallier in collaboration with Marie-Aude Aufaure (Ecole Centrale de Paris), was published in 2012 [20] related to a method for the construction and the automatic settlement of hierarchical structures of concepts. We were particularly interested in the construction of a hierarchical structure of services offered in Hotels from a data set of an application in the field of e-tourism motivated by our contacts with the SME Addictrip. The goal is to associate to each service a concept that provides a common representation of all services. Our experiments are carried out using resources from partners specialized in online hotel booking, in particular from Addictrip. The establishment of a structure of concepts is essential to these partners that use their own terminologies description of hotel services. Indeed it provides a common representation space allowing the comparison of service coming from different resources. Our approach is based on proximity of literal terms in the service having a nearby measure based on n-grams of characters. The results during our experiments show the quality of this approach and its limitations.

5.3. Information and Social Networks Mining for Supporting Information Retrieval

5.3.1. Clustering of Relational Data and Social Network Data

Participants: Yves Lechevallier, Amine Louati, Bruno Almeida Pimentel.

The automatic detection of communities in a social network can provide a kind of graph aggregation. The objective of graph aggregations is to produce small and understandable summaries and it can highlight communities in the network, which greatly facilitates the interpretation.

Social networks allow having a global view of the different actors and different interactions between them, thus facilitating the analysis and information retrieval.

In the enterprise context, a considerable amount of information is stored in relational databases. Therefore, relational database can be a rich source to extract social network.

During this year many updates of the program developed by Louati Amine in 2011 were performed by Bruno Almeida Pimentel. A book chapter, included the new aggregation criteria proposed ans evaluted by Bruno Almeida Pimentel, was written and will be published in 2013.

This work is done in collaboration with Marie-Aude Aufaure, head of the Business Intelligence Team, Ecole Centrale Paris, MAS Laboratory.

5.3.2. Multi-View Clustering on Relational Data

Participants: Thierry Despeyroux, Yves Lechevallier.

In the work reported in [23] in collaboration with Francisco de A.T. de Carvalho, we introduce an improvement of a clustering algorithm described in [17] that is able to partition objects taking into account simultaneously their relational descriptions given by multiple dissimilarity matrices. In this version of the prototype clusters depend on the variables of the representation space. These matrices could have been generated using different sets of variables and dissimilarity functions. This method, which is based on the dynamic clustering algorithm for relational data, is designed to provided a partition and a vector of prototypes for each cluster as well as to learn a relevance weight for each dissimilarity matrix by optimizing an adequacy criterion that measures the fit between clusters and their representatives. These relevance weights change at each algorithm iteration and are different from one cluster to another. Moreover, various tools for the partition and cluster interpretation furnished by this new algorithm are also presented.

Two experiments demonstrate the usefulness of this clustering method and the merit of the partition and cluster interpretation tools. The first one uses a data set from UCI machine learning repository concerning handwritten numbers (digitalized pictures). The second uses a set of reports for which we have an expert classification given a priori. which we have an expert classification given a priori.

5.4. Extension of Usability Methods and Tools

5.4.1. User Evaluation and Tailoring of Personal Information

Participants: Claudia Detraux, Dominique Scapin.

In the context of the ANR project PIMI (Personal Information Management through Internet see section 6.2.1), ergonomic inspections have been carried out to evaluate the usability of the PIMI V0.1. prototype, in its PC and mobile versions [49], [48]. Also, an experiment [24], [35] was conducted on a mockup of a Personal Information Space. Users were asked to perform scenario-based data entry and retrieval tasks, then to modify the mockup according to their wishes and needs. The results allowed to validate the item content and structure for the future personal space, as well as to assess the role of user modifications as evaluation cues, and for the development of further ergonomic recommendations. Detailed information was obtained on how users enter and retrieve data, by modifying the interface settings and shows that the adaptable nature of a Personal Information Space can indeed influence its acceptance, and provides useful cues for ergonomic evaluation

5.4.2. Usability Methods for Information Visualization

Participant: Dominique Scapin.

A collaboration between UFRGS (Federal University of Rio Grande do Sul, Institute of Informatics), Brazil and Inria-AxIS led to a book chapter [37] dealing with potential methodologies for including a user-centered approach into information visualization techniques. It starts by presenting the evolution of visualization techniques evaluation, briefly summarizing the main contributions in this area since its humble beginning as a collateral activity until the recent growth of interest. Then, the focus is on current issues related to such evaluations, particularly concerning the way they are designed and conducted, taking into account a background of well-known usability evaluation methods from HCI to help understanding why there are still open problems. A set of guidelines for a (more) user-centered usability evaluation of information visualization techniques is proposed and discussed.

5.4.3. Usability Recommendations for MIS (Mixed Interactive Systems)

Participant: Dominique Scapin.

A collaboration between University of Toulouse – IUT Tarbes, IRIT and Inria-AxIS led to a book chapter [38] dealing with Mixed Interactive Systems (MIS) which denote an advanced form of interaction that aims at combining physical and digital worlds, such as mixed and augmented reality, tangible user interfaces, ubiquitous computing, etc. Their main interest relates to the use of physical artifacts from the user's activity customary context. The book chapter first reports on a systematic review of the literature on MIS evaluation. From that review, usability recommendations were selected and deciphered before reformulating them under a common format. Finally, three different classification schemes of the usability recommendations obtained are proposed to facilitate search and retrieval, but also to better integrate them into the MIS development process.

5.5. Designing and Evaluating User Experience and Methods for Open Innovation

5.5.1. From Usability to User Experience: an HCI Review

Participants: Dominique Scapin, Bernard Senach, Brigitte Trousse, Marc Pallot.

Through an extensive review of the literature, a paper [28] attempted to characterize a rather novel and popular view on human-computer interaction: User Experience (UX). After introducing its polysemous nature, this paper describes the origins of UX, its scope, underlying concepts and components, as well as its various definitions. Then, UX methods are surveyed and classified, distinguishing processes, frameworks, and specific methods. The paper identifies a set of issues about the needs for increased UX maturity. Even though UX can still be viewed as an extension of usability, its future may correspond to a paradigm evolution rather than simply a buzz word. The evolution is not drastic, but it adds complexity (including new measurements) by considering a few more user areas than traditional usability.

5.5.2. Evaluation of our Methods for Idea Generation Process

Participants: Anne-Laure Negri, Caroline Tiffon, Brigitte Trousse, Bernard Senach.

In 2011 we proposed a methodology coupling two methods [25] (GenIoT a generative method based on probes (fake sensors and/or actuators) and ALoHa! a bodystorming method for designing service concepts in the specific paradigm of the Internet of Things (IoT). In the frame of the European project ELLIOT - Experiential Living Lab for Internet Of Things -, ICT Usage Lab (cf. section 6.1.8) aims at co-creating "green" services, i.e. services based on air quality and noise measurement.

Both IoT ideation methods Aloha! and GenIoT were used for the co-creation of health related services (cf. section 6.3.1.1). The participants of the methods were Environment and Health professional. Results were very different than the workshops run with citizen in the frame of the mobility scenario in ELLIOT (see 2011 AxIS activity report). Comparison of these workshops shows that hybrid approaches –i.e. co-creation approaches mixing both real and virtual meetings are not working as well as pure face to face or pure online approaches. Moreover, GenIoT method seems to be more effective with citizen than with professional. Aloha! is effective in both groups but more efficient with professionals. However the participant experience of Aloha! is higher in the case of citizen (mainly because professionals are not used to practice creative thinking methods and do not appreciate to go out of their comfort zone).

5.5.3. Leading People Behavior Changes: Mining Evolutive Data

Participants: Brigitte Trousse, Yves Lechevallier, Guillaume Pilot, Carole Goffart, Bernard Senach.

The ECOFFICES project (cf. [62], [22] and section 6.1.3) was for AxIS project team our first step towards eco-behavior study. It provided us a very rich context to study how to analyse the evolution of the energy consuming of employees during an energy challenge. A qualitative analysis from questionnaires (before and after the challenge) has been done as well quantitative analysis. The data set for quantitative data is composed of heterogeneous data issued from around 400 sensors (temperature, presence, behavior in terms of opening doors, windows, bonus, malus, etc.). We made different studies related to data preprocessing and data analysis. In our first study [64], we cleaned the data set and selected reliable data for data analysis (only temperature of various equipments, user presence and bonus/malus points). We decided not to work with aggregated variables such as the initial ponderation (defined by partners) for the various bonus-malus rules and the energy consuming at the office level. We decided to use (office, day) as statistical unit (i.e. 9995 units) with a vectoral representation. Finally we realized that the three initial periods (before the challenge, during and after the challenge) on 379 days (2011-2012) should be in fact decomposed in five periods, due to the fact the first and the last periods were split into two subperiods (with and without heat). For the analysis, we apply for each (office,day) a first analysis on a vectorial representation of temperature with the MND method (cf. section focuslab) in order to identify the best partition of these. The MND method uses euclidean distance between each value of the vectorial representation and the prototypes are defined by the means. Second we did a clustering of these units based on bonus and malus and finally we made the correspondence between these two partitions. Three classes for (office, day) are obtained. The interpretation in terms of team relied difficult but we proposed various conclusions for a winner for managing a specific bonus, or in managing ambient temperature or in behavioral change.
In our second study [34] in collaboration with Francisco de A.T. De Carvalho, our goal was to improve the interpretation task at the office and team level by applying AxIS advanced methods. To do this, we applied our hard clustering method presented in [34] on this dataset where each office was characterized by two different representations:

- Interval representation: each office is characterized by a vector of intervals corresponding to the average, minimum and maximum of daily temperatures on the three temperature sensors during these five periods. Then the office is represented by a vector of 15 intervals and the distance used is Hausdorff distance. This classification is consistent with the partition into three classes obtained during the ECOFFICES project. The class obtained with nine eco-responsible ecoffices is the same. However, other offices are divided into two classes according to the type of heating used during the winter period. The classical method divided these offices into two clusters, one of which contains the offices using the radiators during the winter period.
- Sequential representation: Each office is characterized by a vector of 9 measures, the min, max and average of daily temperatures of the three sensors in these five periods. The values are ordinate versus time and the distance used is Urso and Vichi distance (adequate for curves). The results of this approach are quite different from the classical approach results. These results required more effort for their interpretation in collaboration by specialists.

5.5.4. Leading People Behavior Changes: a Literature Review

Participants: Bernard Senach, Anne-Laure Negri.

Our research towards eco-behavior study started with the ECOFFICES project (cf. sections 5.5.3 and 6.1.3 for more details) was recently complemented with a litterature review aiming at a deeper understanding of breaks and levers to eco behavior adoption. A first work was focused on the so-called "modal change problem", compiling methods and tools aimed at supporting people to use public transportation system rather than their personal car. A second work was initiated to get a better understanding of the role that users interface could play in encouraging people to adopt a specific behavior. This work is still in progress.
Eco mobility : prompting people to adopt public transportation mode rather than their personal car.

The first review of work conducted in the fields of Persuasive Communication, Commitment, Nudges and Persuasive Technology showed that behavioral change is a process with many steps requiring to support each step with specific means. For instance, if mass communication can support the public awareness of a problem, information is not sufficient to convince people to really change their behavior. It is necessary to push them to act and numerous well-known influence techniques are nowadays available. All recent technological development (geo localization, mobile devices, social networks) can provide very effective support for behavioral changes as far as they rely on design principles identified by research in Persuasive Technology. A presentation was done on this topic for GreenCode Forum [67] (see the video on youtube).

5.5.5. *Future of Internet and User-Open Innovation for Smart Cities*

Participants: Marc Pallot, Brigitte Trousse, Bernard Senach.

We pursued our work on this topic and contributed to a white paper [59] which is one of the main outcomes of the FIREBALL project [cf. section 6.3.1.2), a Coordination Action within the 7th Framework Programme for ICT, running in the period 2010-2012. The aim of this project was to bring together communities and stakeholders who are active in three areas, namely: research and experimentation on the Future Internet (FIRE); open and user-driven innovation in Living Labs; and urban development for smarter cities. The goal was to develop a common vision on how the different methodologies and concepts in these areas can be aligned for cities as playgrounds of open and user driven innovation related to the Future Internet.

The white paper addresses several aspects that are critical for understanding the 'smart city' concept and the current progress in this area. Based on cases studies and foresight reports we aim to shed light on how the concept of smart city is currently adopted by European Cities and what the ambitions and expectations are in using this concept. It investigates the drivers and bottlenecks that influence the transformation towards a "smart city". Underlying approaches to smart cities are discussed, both in terms of the strategies and planning approaches. From this point of view, this paper explores the conditions that must be established to stimulate the transformation towards smart cities, and the resources that are available or should be made available such as investments in broadband networks and in smart applications, as well as in the capabilities to innovate. This also points to the changing structures and processes of innovation and city development. Interestingly, we see a tendency towards more decentralized and bottom-up approaches to planning and innovation. Innovation ecosystems are characterized by a combination of top down and bottom up initiatives, leading to networking and collaboration among stakeholders, which eventually extend to real innovation communities. Increasingly, citizens, advanced businesses and local governments act as proactive catalysers of innovation, shaping cities as "agents of change".

5.6. FocusLab Platform

5.6.1. *FocusLab platform: software part*

Participants: Brigitte Trousse, Yves Lechevallier, Semi Gaieb, Xavier Augros, Guillaume Pilot, Florian Bonacina.

FocusLab v1.3 (software component) done inside the ELLIOT project (cf. section 6.3.1.1) and for the purposes of the CPER Telius (cf. section 6.1.5) corresponds to the design and the implementation of a set of web-services providing basic and advanced functionalities for data analysis and some other tools supporting the living lab process.

In this version, five data analysis web services are proposed including three generic web services: a classical linear regression and two AxIs methods:

- SMDS/SCDS [91]: SCDS (Sequence Clustering in Data Stream) is a clustering algorithm for mining sequential patterns (Java) in data streams developed by A. Marascu during her thesis. This software takes batches of data in the format "Client-Date-Item" and provides clusters of sequences and their centroids in the form of an approximate sequential pattern calculated with an alignment technique. We propose in this version to return the apparition frequency (min, max, average, slope) of a sequential pattern from data streams (SCDS algorithm) (see references

- GEAR for data streams compression [93], [91], [92], [94]: GEAR (REGLO in french) is an implementation of the history management strategy proposed in Marascu's thesis [1]. It takes a set of time series and provides a memory representation of these series based on a new principle, where salient events are important (in contrast to the recent events of decaying models) .

Other data analysis services and tools have been added for Living Labs needs. We propose also two clustering methods which must be downloaded as standalone software and used for mining data from living labs:

- ATWUEDA (Axis Tool for Web Usage Evolving Data Analysis) for Analysing Evolving Web Usage Data (Da Silva 'thesis 2009 [79], [83], [81], [82]) was developed in Java and uses the JRI library (http://www.r-project.org/). The ATWUEDA tool is able to read data from a cross table in a MySQL database, split the data according to the user specifications (in logical or temporal windows) and then apply the proposed approach in order to detect changes in dynamic environment. Such an approach characterizes the changes undergone by the usage groups (e.g. appearance, disappearance, fusion and split) at each timestamp. Graphics are generated for each analysed window, exhibiting statistics that characterizes changing points over time. This application for the next experiment of Green services use case is under study.
- MND method (Dynamic Clustering Method for Multi-Nominal Data) [90]: The proposed MND method (developed in C++ language) determines iteratively a series of partitions which improves at each step the underlying clustering criterion. The algorithm is based on: a) Prototypes for representing the classes; b) Representation space; c) Proximities (distances or similarities) between two individuals; d) Context-dependent proximity functions for assigning the individuals to the classes at each step. The clustering criterion to be optimized is based on the sum of proximities between individuals and the prototype of the assigning clusters.

  This method has been also successfully applied on Web logs in 2003. This year we improved our code and tested it on IoT data (temperature) issued from the ECOFFICES project (cf. sections 5.5.3 and 6.1.3).

The application of the services provided by FocusLab 1.3 and other AxIS data mining methods for the purposes of ELLIOT use cases and other experimental projects are under study.

IMARA Project-Team

6. New Results

6.1. Low speed automation

Participants: Paulo Lopes Resende, Fawzi Nashashibi, Hao Li, Evangeline Pollard.

The ABV project builds on the HAVEit philosophy (a previous IMARA project for high speed automation) of offering higher levels of automation on highways and organizing the cooperation between human and system along novel automation levels. It differs from HAVEit by focusing on congested traffic at speeds below 50 km/h and adding fully automated driving to the automation spectrum. By automatically following congested traffic, the ABV system relieves the human driver from monotonous tasks. During fully automated driving, the human driver is not required to monitor the system, but has to take over control at the end of the application zone.

6.2. Urban autonomous driving: dealing with intersections

Participants: Guillaume Tréhard, Evangeline Pollard, Fawzi Nashashibi.

The goal of this project, made in collaboration with Valeo is to develop a complete solution for autonomous driving on open roads. More specifically, IMARA's objectives are to provide the way to safely cross any kind of intersections for an autonomous vehicle in a urban context. Among the different relevant scenarios, we can notice:

- Intersection with different shapes: Roundabout, T junctions , X junctions;
- Intersection with different rules: With specific rules (traffic lights, main road...) or unspecified rules ("priority to the right");
- Different traffic: Busy or empty intersections;
- Deal with abnormal situations: road works, policemen, firemen,...

Possible steps for this work can be listed as follows:

- Model the intersection: define relevant information, find a generic model for every intersection;
- Detect the intersection (shape, drivable area, traffic flows);
- Understand the priorities that rules it;
- Locate the car in the intersection by crossing it;
- Plan a path to get out of the intersection.

6.3. Conception of a new communicative system for the protection of vulnerable people

Participants: Pierre Merdrignac, Evangeline Pollard, Oyunchimeg Shagdar, Fawzi Nashashibi.

A new research has been recently launched at IMARA team. The goal is to elaborate a new communicative system between vulnerable people (pedestrian, person with reduced mobility, bicyclist, etc.) and intelligent vehicles in order to improve safety and to limit collision risk. The main idea of this project is as follows. Intelligent vehicles are equipped with an obstacle detection/classification /tracking module in order to prevent injuries. On the other hand, to help the driver in this challenging task, vulnerable people use an application on their mobile phone to inform/share their status on location, type, and dynamics. The status information is transmitted to the driver utilizing wireless communications technology (e.g., 3G and Wi-Fi). In the vehicle, information coming from the communications device and obstacle detection module will be merged to improve the detection and classification tasks. In case of emergency, the vehicle can broadcast safety information to vulnerable people.

6.4. Visible Light V2V Communications for Platooning Control

Participants: Mohammad Abu Alhoul, Oyunchimeg Shagdar, Mohamed Marouf, Fawzi Nashashibi.

Fully automated vehicles have the potential to greatly improve the comfort of humans' life. For driving from one place to another, an automated vehicle must avoid collisions and be able to select non-congested roads for safe and efficient driving. In order to do that the vehicle needs to control its mobility in both macroscopic and microscopic levels by utilizing information exchange with other vehicles and roadside infrastructures based on wireless communications technology. While radio frequency channel is a convincing choose for vehicular communications due to its high data rate over relatively long coverage range (minimum several 100's meters), it is expected to experience channel congestion and low communication reliability especially for the scenario where there is high-density vehicles. In such scenarios vehicles still require to control the mobility on both the macroscopic and microscopic levels, we need to look for supportive and at the same time practical communication media with the ability to support sufficient connection between vehicles. According to the latest standard from IEEE, 802.15.7 for 2011, the communication coverage of Visible Light Communication (VLC) can reach up to 10's of meters, which match the information exchange requirements for mobility control in microscopic level. Motivated by this we started our research activity on modeling of visible light communications channel and design of microscopic mobility control, specifically platooning control, using VLC.

6.5. Augmented reality for the protection of vulnerable people

Participants: Hao Li, Fawzi Nashashibi.

A brand new idea of cooperative augmented reality is under development in IMARA team. It utilizes the results of cooperative local mapping to realize certain augmented reality effect. More specifically, the idea is to obtain an augmented effect of "seeing" through front vehicle, based on the intelligent vehicle sensor configurations.

Given a scenario of two vehicles: a front (first) vehicle and a following (second) vehicle. This front-following vehicles scenario is typical in traffic environment and is potentially dangerous, especially in some occasions such as during an overtaking, where the front vehicle occludes a part of the scene to the following vehicle. The idea of cooperative augmented reality is thus to project the visual perception of the front vehicle onto that of the following vehicle, abiding by perspective geometry. In other words, we patch the occluded part of the view of the following vehicle with corresponding part of the view of the front vehicle. This is not simply a process of partial view copying and pasting between the two vehicles; we have to transform the partial view of the front vehicle according to perspective geometry, in order to make a vivid and natural reproduction of this partial view for the following vehicle, just like if the following vehicle can directly see into the occluded area.

A prerequisite for performing the perspective transformation between the visual perceptions of the two vehicles is the knowledge of the visual perception depth. This knowledge can be estimated by stereo-vision, if correct correspondence is established (yet a challenging process) between the images pair in stereo-vision. However, approximate estimate of the visual perception depth was obtained with the help of 2D range perception in an innovative way and indirect vehicle-to-vehicle relative pose estimation method introduced in [36].

6.6. Step detection for Personal Mobility Vehicles

Participants: Evangeline Pollard, Joshué Pérez Rastelli, Fawzi Nashashibi.

Personal Mobility Vehicles (PMV) is an important part of the Intelligent Transportation System (ITS) domain. These new transport systems have been designed for urban traffic areas, pedestrian streets, green zones and private parks. In these areas, steps and curbs make the movement of disable or mobility reduced people with PMV, and with standard chair wheels difficult. In this work, a step and curb detection system based on laser sensors has been developed. This system is dedicated to vehicles able to cross over steps, for transportation systems, as well as for mobile robots. The system is composed of three laser range finders. Hokuyo UTM 30 LX devices were chosen for their large detection angle (270°) and their high angular resolution (0.25°) and range (30m).

Two laser sensors dedicated to the step detection have a vertical orientation in order to scan the altitude profile of the environment over two lines of sight and the third one, with a vertical orientation is dedicated to obstacle detection.

The step detection process is thus based on the study of the first derivative of the altitude and highlights the use of a new algebraic derivative method (Alien) adapted to laser sensor data. The system has been tested on several real scenarios. It provides the distance, altitude and orientation of the steps in front of the vehicle and offers a high level of precision, even with small steps.

6.7. PROSIVIC-CTS simulator

Participant: Joshué Pérez Rastelli.

The Architecture validation and experiments presented in this document have been implemented in a simulated environment, called ProSiVIC, which allows implementing a virtual Cybercars, among other vehicles. The algorithms are the same as in our Cybus platform, using the position from the SLAM and DGPS sensors. The ProSivic simulator offers a multi-sensorial environment, and takes into account several parameters of a real car such as the inertia, steering wheel response, lateral acceleration with yaw angles, damping suspension, simple weather conditions, friction parameters and more.

Moreover, synchronized time, acceleration (in wheel torque), steering, odometer information, lidar information and camera viewports are some of the components supporting the connection between the control architecture in RTMaps and the simulation.

The simulations show the behavior of the control architecture implemented for CTSs. Two urban scenarios were tested: roundabouts and intersections.

6.8. Autonomous docking based on infrared system for electric vehicle charging in urban areas

Participants: Benjamin Lefaudeux, Joshué Pérez Rastelli, Fawzi Nashashibi.

One of the recent aims of the Intelligent Transportation Systems (ITS) is the reduction of air pollution, reducing the fuel consumption in urban areas and improving road security. To this purpose, electric vehicles are a good and high demanded alternative. Nowadays, some big cities are launching the first electric car-sharing projects to clear its traffic jam, as an alternative to the classic public transportation systems. However, there are still some problems related to energy storage, charging and autonomy to be solved. To tackle this problem in the context of the French project AMARE, IMARA has developed an autonomous docking system, based on an infrared camera embarked in a vehicle equipped with dedicated ADAS, and some infrared diodes installed in the infrastructure, for recharging the vehicle batteries in a street parking area. The results obtained show a good behavior of the implemented system, which is working in a real scenario in the city of Paris.

Different experiments, departing from different points, show a good behavior of the proposed systems. Both lateral and longitudinal errors are lower that the limits of the charging station. The controller used is easy and intuitive for tuning, and the gains can be adapted according of the different vehicles characteristics. This technology permits to assist to human drivers in the charging process of electric vehicles in cities.

6.9. Reasoning for relaxing traffic regulations

Participants: Philippe Morignot, Fawzi Nashashibi.

This work [39] deals with relaxation of traffic rules in unusual but practical situations. For example, if a truck is unloading on a roadway, the automated vehicle should overtake it despite a continuous yellow line: traffic rules are indeed broken, which is illegal, but this might be tolerated due to the unusual aspect of the situation at hand.

An ontology has been developed in order to represent the road network (a directed graph, vertices being intersections and edges being lanes), the infrastructure (road signs, marks), the other road users and the intelligent vehicle. Reasoning on this representation is performed by inference rules (IF/THEN symbolic structures), encoding the deliberation on the encoded situation. Main rules conclude on the next discrete motion of the vehicle, e.g., "pass onto the adjacent lane" which involves crossing a continuous yellow line.

In practice, this ontology has been created using the PROTEGE ontology editor from Stanford University. IF/THEN rules are represented in SWRL (Semantic Web Rule Language), using the reasoner PELLET from the company Clark & Parsia (a plug-in of the tool PROTEGE).

Work over the next year involves porting this reasoning module on the vehicles: porting the generated JAVA source code as one component inside the RTMAPS architecture of CyberCars, and linking the ABoxes (assertional boxes) of the ontology to symbols extracted from signals by perception.

6.10. Communications and Management Control for Cooperative Vehicular Systems

Participants: Ines Ben Jemaa, Oyunchimeg Shagdar, Fawzi Nashashibi, Arnaud de La Fortelle.

One of the attractive applications of electric autonomous vehicles is electric automated Car Sharing service. In this application, a user requests a vehicle at a given geographical location triggering the car sharing system to allocate an autonomous vehicle for the user transport from the station to the user's desired destination. The application requires efficient cooperation among the autonomous vehicles and a service management centre for reliable and responsive car sharing service. Such cooperation is not possible unless vehicles exchange their information on e.g., position, motion, and coordination messages among themselves and with central management entities. While the existing wireless communications technologies can be applied for vehicle to vehicle and vehicle to the infrastructure communications, important research challenges remain including network partitioning problem caused by vehicles' mobility and inability of the convergence of geographically scoped V2V and Internet-based V2I communications. Targeting these issues, we study a topology control solution to tackle the network portioning problem and design of unicast/multicast/Geonetworking schemes for convergence of V2V and V2I communications systems for car-sharing applications [28].

6.11. New urban transportation platforms: Inria's Cybus

Participants: Laurent Bouraoui, François Charlot, Fawzi Nashashibi, Paulo Lopes Resende, Michel Parent, Armand Yvet.

Cybus is the newest prototyping and demonstration platform designed at Inria. Apart from the chassis and engines, the whole hardware and software systems were developed thanks to IMARA's researchers and engineers talents. These electric vehicles are based on a Yamaha chassis but the embedded intelligence is the result of two years of development. Much of the perception and control software has been improved. New guidance functionalities were developed this year, mainly with the introduction of stereovision-based SLAM.

The platforms developed here (Cybus) will be demonstrated in the context of the EU CityMobil-2 project. This time real operational mobility services demonstrations will be extended to 6-12 months in selected European cities! Other showcases are expected to take place in Asian cities in 2014.

6.12. Belief propagation inference for traffic prediction

Participants: Cyril Furtlehner, Yufei Han, Jean-Marc Lasgouttes, Victorin Martin.

This work [57] 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.

Figure 1. The Cybus operated at La Rochelle City during 3 months as a free transport service.

These studies are done in particular in the framework of the projects Travesti and Pumas.

This year's highlights are

- A paper describing a new sufficient condition for local stability of the Belief Propagation algorithm has been published and presented in an international conference [38].
- The work about the theoretical aspects of encoding real valued variables into a binary Ising model has been summarized in a publication currently under reviewing process.
- Ideas about adding macroscopic variables within the Ising model are currently being tested using the software BPstruction developed last year.
- Victorin Martin has given a talk at the "Colloque Jeunes Probabilistes et Statisticiens" at CIRM, where he presented his work one the design of a latent Ising model for real valued inference.
- Cyril Furtlehner, Yufei Han and Victorin Martin presented the work done in the Travesti project at the workshop on inference organized by Inria and Mines ParisTech (see 9.1).

6.13. Non-negative Tensor factorization for spatio-temporal data analysis

Participant: Yufei Han.

This is a joint work with Fabien Moutarde from Mines ParisTech.

We investigate the use of non-negative tensor factorization for spatio-temporal data clustering and prediction. In general case, a spatio-temporal signal is represented as a set of multiple-variant temporal sequences. In the domain of intelligent traffic, the temporal records of traffic flow states (free-flowing/congestion) over a specific time duration with respect to hundreds of links in a transportation network can be considered as a simple but direct example of spatio-temporal signal. Both temporal causality between neighboring time sampling steps and spatial layout of the multiple-variant observation captured at each time sampling step are the focus of the spatio-temporal data analysis. Non-negative tensor factorization enables us to project the high dimensional spatio-temporal data into low-dimensional subspace and clustering/prediction can be then achieved on the derived subspace projection easily.

This year's highlights are

- A conference paper describing application of non-negative tensor factorization in traffic flow state prediction and clustering has been published and presented at ITS World Congress [30];
- The application of non-negative matrix factorization in clustering network-level traffic flow state in large-scale transportation network has been accepted for publication in a journal [11].

6.14. Sparse covariance inverse estimate for Gaussian Markov Random Field

Participants: Cyril Furtlehner, Yufei Han, Jean-Marc Lasgouttes, Victorin Martin.

We investigate in [53] different ways of generating approximate solutions to the inverse problem of pairwise Markov random field (MRF) model learning. We focus mainly on the inverse Ising problem, but discuss also the somewhat related inverse Gaussian problem. In both cases, the belief propagation algorithm can be used in closed form to perform inference tasks. We propose a novel and efficient iterative proportional scaling (IPS) based graph edit method to identify sparse graph linkage of GMRF model to fit underlined data distribution. We remark indeed that both the natural gradient and the best link to be added to a maximum spanning tree solution can be computed analytically. These observations open the way to many possible algorithms, able to find approximate sparse solutions compatible with belief propagation inference procedures and sufficiently flexible to incorporate various spectral constraints like e.g. walk summability. Experimental tests on various data sets with refined L_0 or L_1 regularization procedures indicate that this approach may be a competitive and useful alternative to existing ones.

The part of this work dedicated to Gaussian Markov Random Field has been submitted to the AISTATS 2013 conference.

6.15. Evaluation of dual mode transport system by event-driven simulation

Participants: Arnaud de La Fortelle, Jean-Marc Lasgouttes, Thomas Liennard.

The European project CATS — City Alternative Transport System — is developing and evaluating a new vehicle system using a single type of vehicle for two different usages: individual use or collective transport. Real experiments will necessarily take place with a limited number of vehicles and stations. Hence there is a need for evaluation using simulations. We have been developing a discrete events simulator for that purpose, based on a previous work done for collective taxis [58].

Our model relies on an adapted events/decision graph that extends previous graphs. The new feature of this model is the way we deal with two modes that can be extended to many other modes. This work therefore shows on a concrete example a method to efficiently merge multiple modes into one model.

This year has seen the overhaul of the simulator implementation, as well as the development of a result visualizer that can replay the simulations on a map and show various statistics.

6.16. Multi-speed exclusion processes

Participants: Cyril Furtlehner, Jean-Marc Lasgouttes.

The slow-to-start mechanism is known to play an important role in the particular shape of the fundamental diagram of traffic and to be associated to hysteresis effects of traffic flow. We study this question in the context of stochastic processes, namely exclusion and queueing processes, by including explicitly an asymmetry between deceleration and acceleration in their formulation. Spatial condensation phenomena and metastability are observed, depending on the level of the aforementioned asymmetry. The relationship between these 2 families of models is analyzed on the ring geometry, to yield a large deviation formulation of the fundamental diagram (FD)

This work has been published in the Journal of Statistical Physics [10].

6.17. 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 the stationary regime of the system 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 completed during this year [51] and has been submitted to a journal for publication.

6.18. Exact asymptotics of random walks in the quarter plane

Participant: Guy Fayolle.

In collaboration with K. Raschel (CNRS, Université F. Rabelais à Tours), we pursued the works initiated in 2011.

The enumeration of planar lattice walks, is a classical topic in combinatorics. For a given set $\mathcal{S}$ of allowed unit jumps (or steps), it is a matter of *counting the number of paths* starting from some point and ending at some arbitrary point in a given time, and possibly restricted to some regions of the plane.

Like in the probabilistic context, a common way of attacking these problems relies on the following analytic approach. Let $f(i,j,k)$ denote the number of paths in $\mathbb{Z}_+^2$ starting from $(0,0)$ and ending at (i,j) at time k. Then the corresponding CGF

$$F(x,y,z) = \sum_{i,j,k\geq 0} f(i,j,k)x^i y^j z^k$$

satisfies the functional equation

$$K(x,y)F(x,y,z) = c(x)F(x,0,z) + \widetilde{c}(y)F(0,y,z) + c_0(x,y),$$

where x, y, z are complex variables, although the time variable z plays somehow the role of a parameter. The question of the type of the associated counting generating functions, that is rational, algebraic, holonomic (solution of a linear differential equation with polynomial coefficients), was solved whenever the group is *finite* (see RA 2010). When the group is infinite, the problem is still largely.

It turns out that the nature of singularities play a deep important role in this classification. Making use of the general and powerful approach proposed in the book [2], the paper [9] has been presented at the 23rd International Conference *AofA 2012* on *Combinatorial and Asymptotic Methods for the Analysis of Algorithms*, Montreal, June 17-22.

6.19. Statistical physics and hydrodynamic limits

Participant: Guy Fayolle.

These last years, having in mind a global project concerning the analysis of complex systems, we did focus on the interplay between discrete and continuous description: in some cases, this recurrent question can be addressed quite rigorously via probabilistic methods (see previous activity reports).

To describe the systems of interest, which are in touch with many application domains, we started from *paradigmatic* elements, namely discrete curves subject to stochastic deformations. Up to some convenient mappings, it appears that most models can be set in terms of interacting exclusion processes, the ultimate goal being to derive *hydrodynamic limits* after proper scalings.

The key ideas can be found in [56], where the basic ASEP system on the torus is the toy model. In this case, the usual sequence of empirical measures, converges in probability to a deterministic measure, which is the unique weak solution of a Cauchy problem.

The Gordian knot is indeed the analysis of a family of specific partial differential operators in infinite dimension. Indeed, the values of functions at given points play here the role of usual variables, their number becoming infinite. The method presents some new theoretical features, involving path integrals, promeasures (as introduced by Bourbaki), variational calculus, and the construction of *generalized measures*. In [56], we present a detailed analysis of the ASEP system on the torus $\mathbb{Z}/N\mathbb{Z}$. Then we claim that most of the arguments a priori for multi-type exclusion processes, and should lead to systems of coupled partial differential equations of Burgers' type. At the moment, this claim is being proved for the famous ABC model, reformulated in terms of the dynamics of a random walk on the triangular lattice.

IMEDIA2 Team

6. New Results

6.1. Feature space modeling

Participants: Vera Bakic, Nozha Boujemaa, Esma Elghoul, Hervé Goëau, Sofiene Mouine, Olfa Mzoughi, Anne Verroust-Blondet, Itheri Yahiaoui.

6.1.1. Spatial relations between salient points on a leaf

Participants: Sofiene Mouine, Itheri Yahiaoui, Anne Verroust-Blondet.

We have introduced a novel method for leaf species identification combining local and shape-based features. Our approach extends the shape context model in two ways:
- First of all, two different sets of points are distinguished when computing the shape contexts: the voting set, i.e. the points used to describe the coarse arrangement of the shape and the computing set containing the points where the shape contexts are computed.
Three shape descriptors are proposed, as illustrated in Figure 1 : SC0 (spatial relations between margin points), which corresponds to the original shape context; SC1 (spatial relations between salient points) where the voting set and the computing set are composed of the salient points of the image and SC2 (spatial relations between salient and margin points) where the voting set contains the margin points and the computing set consists of the salient points (see [11] for more details).

Figure 1. From left to right: points used in SC0, SC1 and SC2. The small circles represent the sample points on the leaf margin. The cross points represent the salient points computed with Harris detector.

-This representation is enriched by introducing local features computed in the neighborhood of the computing points.
We obtained excellent identification scores in the ImageCLEF 2012 plant identification task for scan and scan-like images of leaves (RUN2 in [20]).

6.1.2. Detection and extraction of leaf parts for plant identification

Participants: Olfa Mzoughi, Itheri Yahiaoui, Nozha Boujemaa.

Automatic plant identification is a relatively new research area in computer vision that has increasingly attracted high interest as a promising solution for the development of many botanical industries and for the success of biodiversity conservation. Most of the approaches proposed are based on the analysis of morphological properties of leaves. They have applied several well-known generic shape descriptors. Nevertheless, faced with the large amount of leaf species, botanical knowledge, especially about leaf parts (petiole, blade, insertion point, base and apex, rachis) is important to enhance their precision.
First of all, in order to extract them from leaf images, we introduced two types of symmetry in [12]: (i) the local translational symmetry, which is useful for petiole and rachis detection and (ii) the local symmetry of depth indentations, which is suited for base and apex detection.
Then, we studied the usefulness of parts detection (mainly petiole and insertion point) as a pre-processing stage for classic leaf shape retrieval schemes [13]. We showed that the removal of the petiole and the use of the insertion point as a starting point for the descriptors sensitive to the starting point improve retrieval results.

6.1.3. Multi-organ plant identification

Participants: Hervé Goëau, Vera Bakic, Souheil Selmi.

Inspired by citizen sciences, the main goal of this work is to speed up the collection and integration of raw botanical observation data, while providing to potential users an easy and efficient access to this botanical knowledge. We therefore designed and developed an original crowd-sourcing web application dedicated to the access of botanical knowledge through automated identification of plant species by visual content with multi-organ queries. Technically, the first side of the application deals with multi-organ content-based identification of plant. Indeed, most methods proposed for such automatic identification are actually based on leaf images, mostly based on leaf segmentation and boundary shape. However, leaves are far from being the only discriminant visual key between species and they are not visible all over the year for a large fraction of plant species. We propose to make the use of five different organs and plant's views including habit, flowers, fruits, leaves and bark. Thanks to an interactive and visual query widget, the tagging process of the different organs and views is as simple as drag-and-drop operations and does not require any expertise in botany.

Technically, as suggested by the results of ImageCLEF2011 for leaves [24], it is based on local features and large-scale matching. Interest points are detected with a modified color Harris detector, in order to favor points with a central position in the image and to reduce the impact of background features. Each interest point is then described with a SURF local feature and an HSV histogram. Automatic system-oriented and human-centered evaluations of the application show that the results are already satisfactory and therefore very promising in the long term to identify a richer flora. The second side of the application deals with interactive tagging and allows any user to validate or correct the automatic determinations returned by the system. Overall, this collaborative system enables the automatic and continuous enrichment of the visual botanical knowledge and therefore it increases progressively the accuracy of the automated identification. This application called 'Identify' (cf. Figure 2 and http://identify.plantnet-project.org) has been presented at the first ACM International workshop on Multimedia Analysis and Ecological Data [8]. This work has been done in collaboration with Inria team ZENITH and with the botanists of the AMAP UMR team (CIRAD). It is also closely related to a citizen science project around plant's identification that we developed with the support of the Tela Botanica social network inside the Pl@ntNet project.

6.1.4. Segmentation transfer method for articulated models

Participants: Esma Elghoul, Anne Verroust-Blondet.

Mesh segmentation consists in partitioning the surface into a set of patches that are uniform with respect to a given property. We are interested in retaining the semantic information during the segmentation. A particularly challenging task is then the automatic identification of semantically meaningful parts of a 3D model, which can be hard to achieve when only the shape geometry is considered. We have introduced a method using a pre-segmented example model to perform semantic-oriented segmentation of non-rigid 3D models of the same class (human, octopus, quadrupeds, etc.). Using the fact that the same type of non-rigid models share the same global topological structure, we exploit coarse topological shape attributes in conjunction with a seed-based segmentation approach to transfer a meaningful and consistent segmentation from the example mesh to the

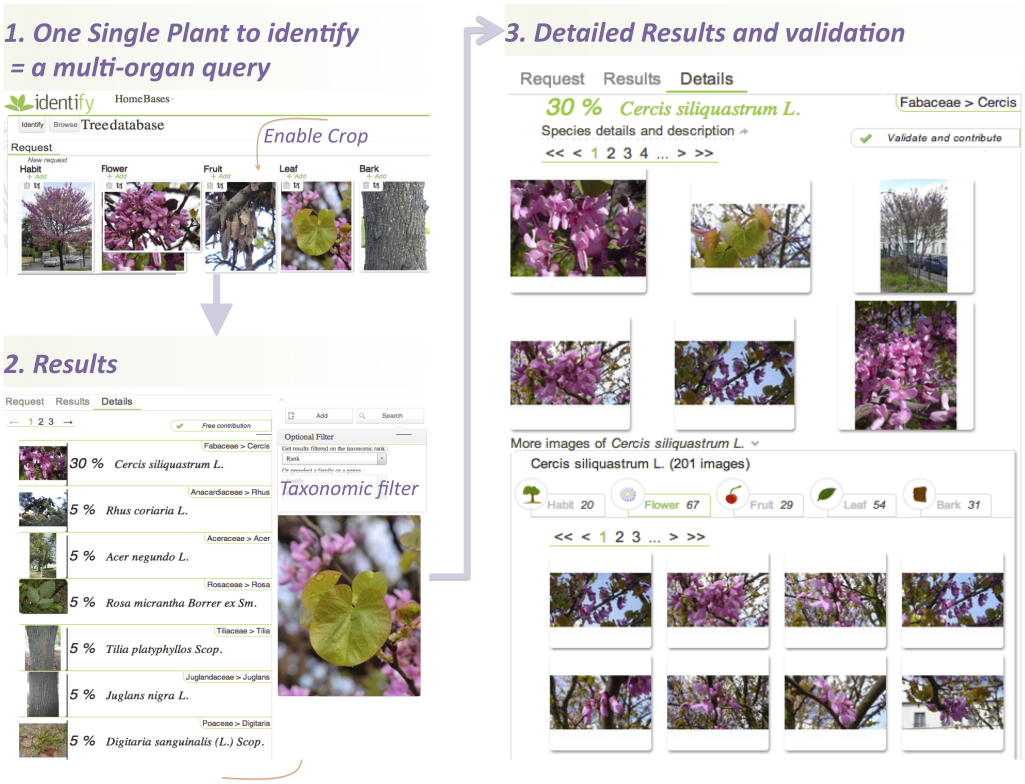

Figure 2. Example of a multi-organ query on one single plant submitted in the application.

target models. Promising results have been obtained on classes of articulated models (cf. Figure 3). This work has been submitted for publication.

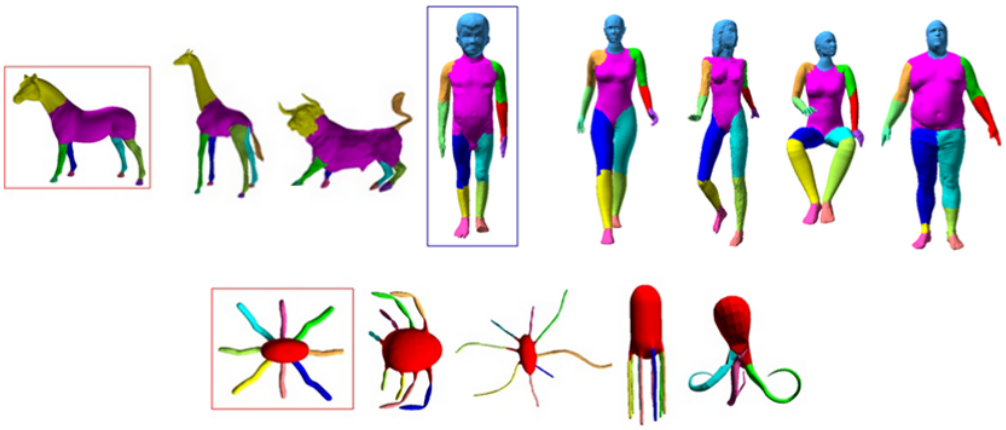

Figure 3. Segmentation transfer results on the quadrupeds, on the humans and on the octopus class. The exemplar segmentations are framed in red or in blue.

6.2. Feature space structuring

Participants: Nozha Boujemaa, Hervé Goëau, Amel Hamzaoui, Saloua Ouertani-Litayem, Mohamed Riadh Trad.

6.2.1. Plant Leaves Morphological Categorization with Shared Nearest Neighbors Clustering

Participants: Amel Hamzaoui, Hervé Goëau, Nozha Boujemaa.

In [9] we present an original experiment aimed at evaluating if state-of-the-art visual clustering techniques are able to automatically recover morphological classifications built by the botanists themselves. The clustering phase is based on a recent Shared-Nearest Neighbors (SNN) clustering algorithm, which allows combining effectively heterogeneous visual information sources at the category level. Each resulting cluster is associated with an optimal selection of visual similarities, allowing discovering diverse and meaningful morphological categories even if we use a blind set of visual sources as input. Experiments have been performed on ImageCLEF 2011 plant identification dataset [23], specifically enriched in this work with morphological attributes tags (annotated by expert botanists). The results presented in Figure 4 are very promising, since all clusters discovered automatically can be easily matched to one node of the morphological tree built by the botanists. This work is also described in details in Amel Hamzaoui's thesis [4].

6.2.2. Distributed KNN-Graph approximation via Hashing

Participants: Mohamed Riadh Trad, Nozha Boujemaa.

High dimensional data hashing is essential for scaling up and distributing data analysis applications involving feature-rich objects, such as text documents, images or multi-modal entities (scientific observations, events, etc.). In this first research track, we first investigated the use of high dimensional hashing methods for efficiently approximating K-NN Graphs [16], [19], [17], particularly in distributed environments. We highlighted the importance of balancing issues on the performance of such approaches and show why the baseline approach

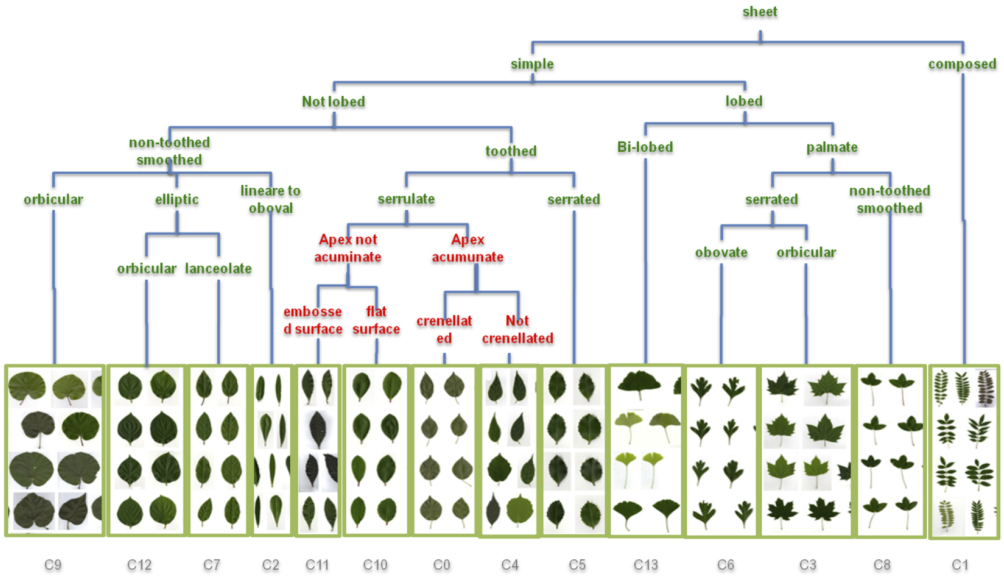

Figure 4. Hierarchical tree organization of the clusters produced by the SNN clustering method on the scan leaf dataset ImageCLEF2011.

using Locality Sensitive Hashing does not perform well. Our new KNN-join method is based on RMMH, a hash function family based on randomly trained classifiers that we introduced in 2011. We show that the resulting hash tables are much more balanced and that the number of resulting collisions can be greatly reduced without degrading quality. We further improve the load balancing of our distributed approach by designing a parallelized local join algorithm, implemented within the MapReduce framework.

6.2.3. Hash-Based Support Vector Machines Approximation for Large Scale Prediction

Participants: Saloua Ouertani-Litayem, Nozha Boujemaa.

How-to train effective classifiers on huge amount of multimedia data is clearly a major challenge that is attracting more and more research works across several communities. Less efforts however are spent on the counterpart scalability issue: how to apply big trained models efficiently on huge non annotated media collections ? In [10], we addressed the problem of speeding-up the prediction phase of linear Support Vector Machines via Locality Sensitive Hashing. We proposed building efficient hash-based classifiers that are applied in a first stage in order to approximate the exact results and filter the hypothesis space. Experiments performed with millions of one-against-one classifiers show that the proposed hash-based classifier can be more than two orders of magnitude faster than the exact classifier with minor losses in quality (cf. Figure 5).

6.3. Pattern recognition and statistical learning

Participants: Nozha Boujemaa, Michel Crucianu, Donald Geman, Wajih Ouertani, Asma Rejeb Sfar.

6.3.1. Machine identification of biological shapes

Participants: Asma Rejeb Sfar, Donald Geman, Nozha Boujemaa.

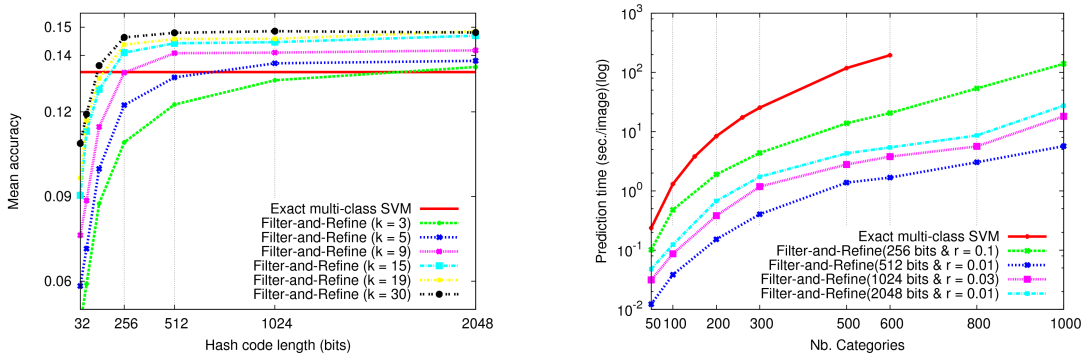

Figure 5. Exact Multi-class SVM vs HBMS based Filter-And-Refine method in terms of accuracy and prediction time

The increasing availability of digital images in the traditional sciences, the growing interest in biodiversity and the ongoing shortage of skilled taxonomists combine to make the automated categorization algorithms, increasingly important in many fields such as botany, agriculture and medicine. In this work, we propose a hierarchical coarse-to-fine approach to identify botanical species from a scanned sample of a plant organ, e.g., a leaf or a flower. To this end, we exploit domain-specific knowledge about taxonomy and landmarks. Promising recognition rates are achieved on several leaf datasets. Results have been submitted for publication.

6.3.2. Relevance feedback on partial image query

Participants: Wajih Ouertani, Michel Crucianu, Nozha Boujemaa.

scalability, hashing, SVM, prediction, approximation

Even if cropping an image to perform one-shot partial query filters a considerable amount of senseless regions for target definition, it does not yet clearly illustrates what the user is looking for. Indeed, the user target is either closer to the instance level or to the category level. Then we may have numerous suggested examples within the first response ranks while possibly some of them are totally irrelevant examples.

We claim that a localization interaction is still more appropriate than having a holistic decision about image relevance if it is performed on more examples. We go beyond the first partial query and investigate machine learning process to learn intention iteratively and interactively. Our learning process is based on what user delimit within additional images taken from the first response ranks. Our motivations include dealing with semantic gap revealed by local features hit falling into false regions within retrieved images. Those images might be either totally irrelevant, where all partial zones are out of the interest, or partially relevant, not because of the zones expected by the system (false-localization) but rather because of some missed zones. Through local annotations we expect the ability of redirecting the recognition session to those relevant regions and studying how much we can reduce the semantic gap within interactive localization.

This year, we studied several learning strategies based on several assumptions heuristically extracted on user interaction. The presented strategies have been also combined with features filtering within object representation. The filtering includes grouping contextualizing and varying features set representations.

6.4. IKONA/MAESTRO sofware

Participants: Vera Bakic, Laurent Joyeux, Sofiene Mouine, Souheil Selmi.

This year, IKONA has been extended in the context of Pl@ntNet, Glocal and I-SEARCH projects.

- For the Pl@ntNet project, along the continuing improvements and optimizations in the MAESTRO software, a number of new features were added:

new options for interest points distribution and filtering with the segmented image,
a new shape context descriptor (corresponding to [11]),
various combinations of descriptors in one vector per interest point or region,
for regions: extraction of sub-images, EOH and Fourier descriptors,
more options for update of calculated signatures, new score type (used in ImageCLEF2012) and decision rules (adaptive Knn calculation based on individual plant information) for statistical tools.
In addition, a number of new web services and functionality were developed/updated and deployed:
the addition of new databases (Vignes, Musa), while some with the organ annotation (Photoflora, Girod), and the update of the multi search views for the new datasets;
the development of new services allowing to return of botanical information in several formats (csv, xml, JSON...);
the update of the indexation system so that it can crawl images from different sources (internal or external sources like CEL web service, which uses html or identiplante web service (in JSON format);
the development of html pages to annotate a set of images by organ type;
the development of an applet demonstration of leaf architecture (this applet interfaces a library developed by a PhD student working on this project);
the development of an application "Pl@ntNet Identify" for android platform and its interfacing with the existing web services.

- For the Glocal project, functionalities such as fraud detection and similarity search were integrated in the mock up of user interface and in the final demonstration of the project. In addition, a dynamic indexation system of images from AFP (Agence France-Presse) was implemented as well as the similarity web services working on this dynamic dataset.
- For the I-SEARCH project, an integration of the video mining component in I-SEARCH platform was done. The component extracts visual objects that are the most recurrent from a set of images, or in a video.

SMIS Project-Team

6. New Results

6.1. Embedded Data Management

Participants: Nicolas Anciaux, Luc Bouganim, Lionel Le Folgoc, Yanli Guo, Saliha Lallali, Philippe Pucheral, Iulian Sandu Popa, Shaoyi Yin.

Inspired by low cost economic models, this work draws the idea of a one-dollar database machine, with the objective to disseminate databases everywhere, up to the lightest smart objects. In contrast to traditional database machines relying on massively parallel architectures, the one-dollar database machine considers the cheapest form of computer available today: a microcontroller equipped with GBs size (external) Flash storage. Designing such a database machine is very challenging due to a combination of conflicting RAM and NAND Flash constraints. To tackle this challenge, this work proposes a new paradigm based on database serialization (managing all database structures in a pure sequential way) and stratification (restructuring them into strata when a scalability limit is reached).We show that a complete DBMS engine can be designed according to this paradigm and demonstrate the effectiveness of the approach through a performance evaluation. This work capitalizes on previous results related to the indexing of Flash resident data [16] and has also obvious connections with the more general study we are conducting on Flash-based data management (see Section 6.2). Partial elements of this solution have been demonstrated at [13]. In 2012, we have extended our previous results on indexation of flash resident data [IS] and we have proposed the design of a complete DBMS engine [DAPD] complying by nature with the conflicting RAM and NAND Flash constraints we are facing. Currently, we work at the extension of the embedded DBMS engine to support document data (e.g., text documents or any type of documents that are tagged)) and spatio-temporal data (e.g., vehicle trajectory data or any type of time-stamped and/or geo-located data).

6.2. Flash-based Data Management

Participants: Matias Bjørling, Philippe Bonnet, Luc Bouganim, Niv Dayan.

Solid State Drives (SSDs) are replacing magnetic disks as secondary storage for database management, as they offer orders of magnitude improvement in terms of bandwidth and latency. In terms of system design, the advent of SSDs raises considerable challenges. First, the storage chips, which are the basic component of a SSD, have widely different characteristics – e.g., copy-on-write, erase-before-write and page-addressability for flash chips vs. in-place update and byte-addressability for PCM chips. Second, SSDs are no longer a bottleneck in terms of I/O latency forcing streamlined execution throughout the I/O stack. Finally, SSDs provide a high degree of parallelism that must be leveraged to reach nominal bandwidth. This evolution puts database system researchers at a crossroad. The first option is to hang on to the current architecture where secondary storage is encapsulated behind a block device interface. This is the mainstream option both in industry and academia. This leaves the storage and OS communities with the responsibility to deal with the complexity introduced by SSDs in the hope that they will provide us with a robust, yet simple, performance model. We showed that this option amounts to building on quicksand. We illustrated our point by debunking some popular myths about flash devices and by pointing out mistakes in the papers we have published throughout the years. The second option is to abandon the simple abstraction of the block device interface and reconsider how database storage managers, operating system drivers and SSD controllers interact. We gave our vision of how modern database systems should interact with secondary storage. This approach requires a deep re-design of the database system architecture, which is the only viable option for database system researchers to avoid becoming irrelevant. This work started at the end of 2011 and was published at CIDR'13 [20], in cooperation with the IT University of Copenhagen.

6.3. Minimal Exposure

Participants: Nicolas Anciaux, Walid Bezza, Danae Boutara, Benjamin Nguyen, Michalis Vazirgiannis.

When users request a service, the service provider usually asks for personal documents to tailor its service to the specific situation of the applicant. For example, the rate and duration of consumer's loans are usually adapted depending on the risk based on the income, assets or past lines of credits of the borrower. In practice, an excessive amount of personal data is collected and stored. Indeed, a paradox is at the root of this problem: service providers require users to expose data in order to determine whether that data is needed or not to achieve the purpose of the service. We explore a reverse approach, where service providers would publicly describe the data they require to complete their task, and where software (placed, depending on the context, on the client, on the server, or in a trusted hardware component) would use those descriptions to determine a minimum subset of information to expose. In 2012, we have presented our general framework called Minimum Exposure [14], we have modelled the underlying problem (for simple tasks) and proposed resolution algorithms [19], [24], and we have addressed the case of multi-label classifiers [18]. In the short term, we plan to adapt the minimum exposure architecture to support hidden decision rules using smart cards. Then, we will investigate new privacy metrics to capture the degree of exposure of sets of personal data items better.

6.4. Secure Global Computing on Asymmetric Architecture

Participants: Tristan Allard, Benjamin Nguyen, Philippe Pucheral, Quoc-Cuong To.

This research direction is based on the 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. Trust is assumed ad hoc and can be justified by the use of secure tokens, open source software, friend relationships etc. In our work, we use tamper resistant secure tokens running trusted software, which provide a high degree of trust, due to the overwhelming cost of hardware tampering. The main difficulty on such an architecture is global processing i.e. constructing aggregate data from the individual records, because the entity in charge of executing the global computation is untrusted. Given our large scale data centric applications (e.g. nationwide surveys), we also discard solutions based on secure multi-party computation, which do not scale. We have studied 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. This work is an extension of [31] We are now studying more generally the execution of SQL "Group by" queries on this architecture, which is the topic of Quoc-Cuong To's Ph.D. thesis started in sept. 2012. We have published preliminary results on this novel problem in [23], which adapts the techniques proposed in [31].

6.5. Trusted Cell Data Management

Participants: Nicolas Anciaux, Philippe Bonnet, Luc Bouganim, Benjamin Nguyen, Philippe Pucheral, Iulian Sandu Popa.

With the convergence of mobile communications, sensors and online social networks technologies, we are witnessing an exponential increase in the creation and consumption of personal data. Such data is volunteered by users, automatically captured by sensors or inferred from existing data. Today, there is a wide consensus that individuals should have increased control on how their personal data is collected, managed and shared. Yet there is no appropriate technical solution to implement such personal data services: centralized solutions sacrifice security for innovative applications, while decentralized solutions sacrifice innovative applications for security. In this work, we argue that the advent of secure hardware in all personal IT devices, at the edges of the Internet, could trigger a sea change. We propose the vision of trusted cells: personal data servers running on secure smart phones, set-top boxes, secure portable tokens or smart cards to form a global, decentralized data platform that provides security yet enables innovative applications. We motivate our approach, describe the trusted cells architecture and define a range of challenges for future research in a paper published at CIDR'13 (Int. Conf on Innovative Data Systems Research) [17].

6.6. Experiment in the medical field

Participants: Nicolas Anciaux, Luc Bouganim, Philippe Pucheral, Alexei Troussov.

The PlugDB engine is being experimented in the field since September 2011 to implement a secure and portable medical-social folder. The objective is to improve the coordination of medical care and social services provided at home for dependent people. Details related to this experiment are given in Section 7.2 . While this action did not generate new academic results (though it helped us validating some previous results), it imposed us a strong investment in terms of test and optimization for our prototype and in terms of communication to promote this experiment at the regional level.

WILLOW Project-Team

6. New Results

6.1. 3D object and scene modeling, analysis, and retrieval

6.1.1. People Watching: Human Actions as a Cue for Single View Geometry

Participants: Vincent Delaitre, Ivan Laptev, Josef Sivic, Alexei Efros [CMU], David Fouhey [CMU], Abhinav Gupta [CMU].

We present an approach which exploits the coupling between human actions and scene geometry. We investigate the use of human pose as a cue for single-view 3D scene understanding. Our method builds upon recent advances in still-image pose estimation to extract functional and geometric constraints about the scene. These constraints are then used to improve state-of-the-art single-view 3D scene understanding approaches. The proposed method is validated on a collection of monocular time lapse sequences collected from YouTube and a dataset of still images of indoor scenes. We demonstrate that observing people performing different actions can significantly improve estimates of 3D scene geometry.

This work has been published in [11].

6.1.2. Learning and Calibrating Per-Location Classifiers for Visual Place Recognition

Participants: Petr Gronát, Josef Sivic, Guillaume Obozinski [Inria SIERRA], Tomáš Pajdla [CTU in Prague].

The aim of this work is to localize a query photograph by finding other images depicting the same place in a large geotagged image database. This is a challenging task due to changes in viewpoint, imaging conditions and the large size of the image database. The contribution of this work is two-fold. First, we cast the place recognition problem as a classification task and use the available geotags to train a classifier for each location in the database in a similar manner to per-exemplar SVMs in object recognition. Second, as only few positive training examples are available for each location, we propose a new approach to calibrate all the per-location SVM classifiers using *only* the negative examples. The calibration we propose relies on a significance measure essentially equivalent to the p-values classically used in statistical hypothesis testing. Experiments are performed on a database of 25,000 geotagged street view images of Pittsburgh and demonstrate improved place recognition accuracy of the proposed approach over the previous work. The problem addressed in this work is illustrated in Figure 1 .

This work has been submitted to CVPR 2013.

6.1.3. What Makes Paris Look like Paris?

Participants: Josef Sivic, Carl Doersch [CMU], Saurabh Singh [UIUC], Abhinav Gupta [CMU], Alexei Efros [CMU].

Given a large repository of geotagged imagery, we seek to automatically find visual elements, e.g. windows, balconies, and street signs, that are most distinctive for a certain geo-spatial area, for example the city of Paris. This is a tremendously difficult task as the visual features distinguishing architectural elements of different places can be very subtle. In addition, we face a hard search problem: given all possible patches in all images, which of them are both frequently occurring and geographically informative? To address these issues, we propose to use a discriminative clustering approach able to take into account the weak geographic supervision. We show that geographically representative image elements can be discovered automatically from Google Street View imagery in a discriminative manner. We demonstrate that these elements are visually interpretable and perceptually geo-informative. The discovered visual elements can also support a variety of computational geography tasks, such as mapping architectural correspondences and influences within and across cities, finding representative elements at different geo-spatial scales, and geographically-informed image retrieval. Example result is shown in Figure 2 .

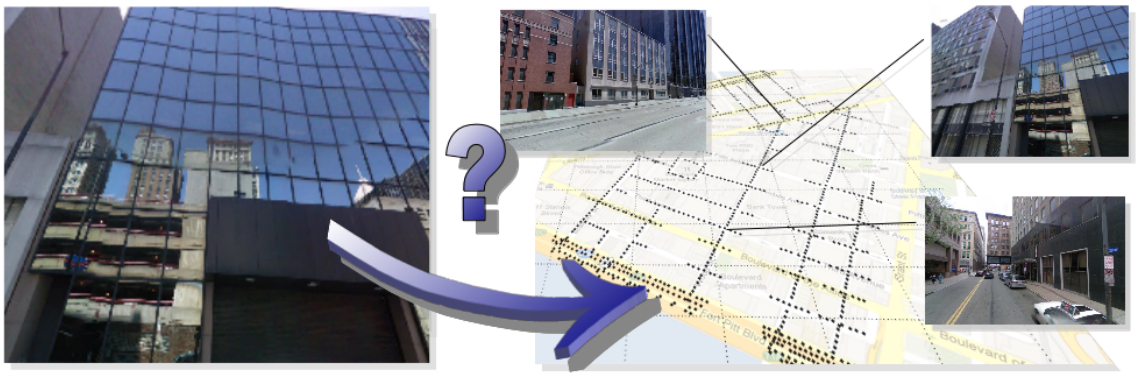

Figure 1. The goal of this work is to localize a query photograph (left) by finding other images of the same place in a large geotagged image database (right). We cast the problem as a classification task and learn a classifier for each location in the database. We develop a non-parametric procedure to calibrate the outputs of the large number of per-location classifiers without the need for additional positive training data.

Figure 2. Examples of geographic patterns in Paris (shown as red dots on the maps) for three discovered visual elements (shown below each map). Balconies with cast-iron railings are concentrated on the main boulevards (left). Windows with railings mostly occur on smaller streets (middle). Arch supporting columns are concentrated on Place des Vosges and the St. Germain market (right).

This work has been published in [6].

6.2. Category-level object and scene recognition

6.2.1. Task-Driven Dictionary Learning

Participants: Jean Ponce, Julien Mairal [Inria LEAR], Francis Bach [Inria SIERRA].

Modeling data with linear combinations of a few elements from a learned dictionary has been the focus of much recent research in machine learning, neuroscience and signal processing. For signals such as natural images that admit such sparse representations, it is now well established that these models are well suited to restoration tasks. In this context, learning the dictionary amounts to solving a large-scale matrix factorization problem, which can be done efficiently with classical optimization tools. The same approach has also been used for learning features from data for other purposes, e.g., image classification, but tuning the dictionary in a supervised way for these tasks has proven to be more difficult. In this paper, we present a general formulation for supervised dictionary learning adapted to a wide variety of tasks, and present an efficient algorithm for solving the corresponding optimization problem. Experiments on handwritten digit classification, digital art identification, nonlinear inverse image problems, and compressed sensing demonstrate that our approach is effective in large-scale settings, and is well suited to supervised and semi-supervised classification, as well as regression tasks for data that admit sparse representations.

This work has been published in [7].

6.2.2. Object Detection Using Strongly-Supervised Deformable Part Models

Participants: Ivan Laptev, Hossein Azizpour [KTH].

Deformable part-based models achieve state-of-the-art performance for object detection, but rely on heuristic initialization during training due to the optimization of non-convex cost function. This work investigates limitations of such an initialization and extends earlier methods using additional supervision. We explore strong supervision in terms of annotated object parts and use it to (i) improve model initialization, (ii) optimize model structure, and (iii) handle partial occlusions. Our method is able to deal with sub-optimal and incomplete annotations of object parts and is shown to benefit from semi-supervised learning setups where part-level annotation is provided for a fraction of positive examples only. Experimental results are reported for the detection of six animal classes in PASCAL VOC 2007 and 2010 datasets. We demonstrate significant improvements in detection performance compared to the LSVM and the Poselet object detectors.

This work has been published in [9].

6.2.3. Multi-Class Cosegmentation

Participants: Armand Joulin, Jean Ponce, Francis Bach [Inria SIERRA].

Bottom-up, fully unsupervised segmentation remains a daunting challenge for computer vision. In the cosegmentation context, on the other hand, the availability of multiple images assumed to contain instances of the same object classes provides a weak form of supervision that can be exploited by discriminative approaches. Unfortunately, most existing algorithms are limited to a very small number of images and/or object classes (typically two of each). This work proposes a novel energy-minimization approach to cosegmentation that can handle multiple classes and a significantly larger number of images. The proposed cost function combines spectral- and discriminative-clustering terms, and it admits a probabilistic interpretation. It is optimized using an efficient EM method, initialized using a convex quadratic approximation of the energy. Comparative experiments show that the proposed approach matches or improves the state of the art on several standard datasets.

This work has been published in [13].

6.2.4. A Convex Relaxation for Weakly Supervised Classifiers

Participants: Armand Joulin, Francis Bach [Inria SIERRA].

This work introduces a general multi-class approach to weakly supervised classification. Inferring the labels and learning the parameters of the model is usually done jointly through a block-coordinate descent algorithm such as expectation-maximization (EM), which may lead to local minima. To avoid this problem, we propose a cost function based on a convex relaxation of the soft-max loss. We then propose an algorithm specifically designed to efficiently solve the corresponding semidefinite program (SDP). Empirically, our method compares favorably to standard ones on different datasets for multiple instance learning and semi-supervised learning, as well as on clustering tasks.

This work has been published in [12].

6.2.5. Top-Down and Bottom-Up Cues for Scene Text Recognition

Participants: Karteek Alahari, Anand Mishra [IIT India], C.V. Jawahar [IIT India].

Scene text recognition has gained significant attention from the computer vision community in recent years. Recognizing such text is a challenging problem, even more so than the recognition of scanned documents. In this work, we focus on the problem of recognizing text extracted from street images. We present a framework that exploits both bottom-up and top-down cues. The bottom-up cues are derived from individual character detections from the image. We build a Conditional Random Field model on these detections to jointly model the strength of the detections and the interactions between them. We impose top-down cues obtained from a lexicon-based prior, i.e. language statistics, on the model. The optimal word represented by the text image is obtained by minimizing the energy function corresponding to the random field model.

We show significant improvements in accuracies on two challenging public datasets, namely Street View Text (over 15%) and ICDAR 2003 (nearly 10%).

This work has been published in [15].

6.2.6. Scene Text Recognition using Higher Order Language Priors

Participants: Karteek Alahari, Anand Mishra [IIT India], C.V. Jawahar [IIT India].

The problem of recognizing text in images taken in the wild has gained significant attention from the computer vision community in recent years. Contrary to recognition of printed documents, recognizing scene text is a challenging problem. We focus on the problem of recognizing text extracted from natural scene images and the web. Significant attempts have been made to address this problem in the recent past. However, many of these works benefit from the availability of strong context, which naturally limits their applicability. In this work we present a framework that uses a higher order prior computed from an English dictionary to recognize a word, which may or may not be a part of the dictionary. We show experimental results on publicly available datasets. Furthermore, we introduce a large challenging word dataset with five thousand words to evaluate various steps of our method exhaustively.

The main contributions of this work are: (1) We present a framework, which incorporates higher order statistical language models to recognize words in an unconstrained manner (i.e. we overcome the need for restricted word lists, and instead use an English dictionary to compute the priors). (2) We achieve significant improvement (more than 20%) in word recognition accuracies without using a restricted word list. (3) We introduce a large word recognition dataset (at least 5 times larger than other public datasets) with character level annotation and benchmark it.

This work has been published in [14].

6.3. Image restoration, manipulation and enhancement

6.3.1. Non-Uniform Deblurring for Shaken Images

Participants: Josef Sivic, Andrew Zisserman, Jean Ponce, Oliver Whyte [Microsoft Redmond].

Photographs taken in low-light conditions are often blurry as a result of camera shake, i.e. a motion of the camera while its shutter is open. Most existing deblurring methods model the observed blurry image as the convolution of a sharp image with a uniform blur kernel. However, we show that blur from camera shake is in general mostly due to the 3D rotation of the camera, resulting in a blur that can be significantly non-uniform across the image. We propose a new parametrized geometric model of the blurring process in terms of the rotational motion of the camera during exposure. This model is able to capture non-uniform blur in an image due to camera shake using a single global descriptor, and can be substituted into existing deblurring algorithms with only small modifications. To demonstrate its effectiveness, we apply this model to two deblurring problems; first, the case where a single blurry image is available, for which we examine both an approximate marginalization approach and a maximum a posteriori approach, and second, the case where a sharp but noisy image of the scene is available in addition to the blurry image. We show that our approach makes it possible to model and remove a wider class of blurs than previous approaches, including uniform blur as a special case, and demonstrate its effectiveness with experiments on synthetic and real images.

This work has been published in [8]. An image deblurring demo, described in section 5.8 , has been made available online.

6.3.2. Learning to Estimate and Remove Non-uniform Image Blur

Participants: Florent Couzinie-Devy, Jian Sun, Karteek Alahari, Jean Ponce.

This work addresses the problem of restoring images subjected to unknown and spatially varying blur caused by defocus or linear (say, horizontal) motion. The estimation of the global (non-uniform) image blur is cast as a multi-label energy minimization problem. The energy is the sum of unary terms corresponding to learned local blur estimators, and binary ones corresponding to blur smoothness. Its global minimum is found using Ishikawa's method by exploiting the natural order of discretized blur values for linear motions and defocus. Once the blur has been estimated, the image is restored using a robust (non-uniform) deblurring algorithm based on sparse regularization with global image statistics. The proposed algorithm outputs both a segmentation of the image into uniform-blur layers and an estimate of the corresponding sharp image. We present qualitative results on real images, and use synthetic data to quantitatively compare our approach to the publicly available implementation of Chakrabarti et al. 2010.

This work has been submitted to CVPR 2013.

6.4. Human activity capture and classification

6.4.1. Scene Semantics from Long-Term Observation of People

Participants: Vincent Delaitre, Ivan Laptev, Josef Sivic, David Fouhey [CMU], Abhinav Gupta [CMU], Alexei Efros [CMU].

Our everyday objects support various tasks and can be used by people for different purposes. While object classification is a widely studied topic in computer vision, recognition of object function, i.e., what people can do with an object and how they do it, is rarely addressed. In this work we construct a functional object description with the aim to recognize objects by the way people interact with them. We describe scene objects (sofas, tables, chairs) by associated human poses and object appearance. Our model is learned discriminatively from automatically estimated body poses in many realistic scenes. In particular, we make use of time-lapse videos from YouTube providing a rich source of common human-object interactions and minimizing the effort of manual object annotation. We show how the models learned from human observations significantly improve object recognition and enable prediction of characteristic human poses in new scenes. Results are shown on a dataset of more than 400,000 frames obtained from 146 time-lapse videos of challenging and realistic indoor scenes. Some of the estimated human poses and results of pixel-wise scene segmentation are shown in Figure 3 .

This work has been published in [10].

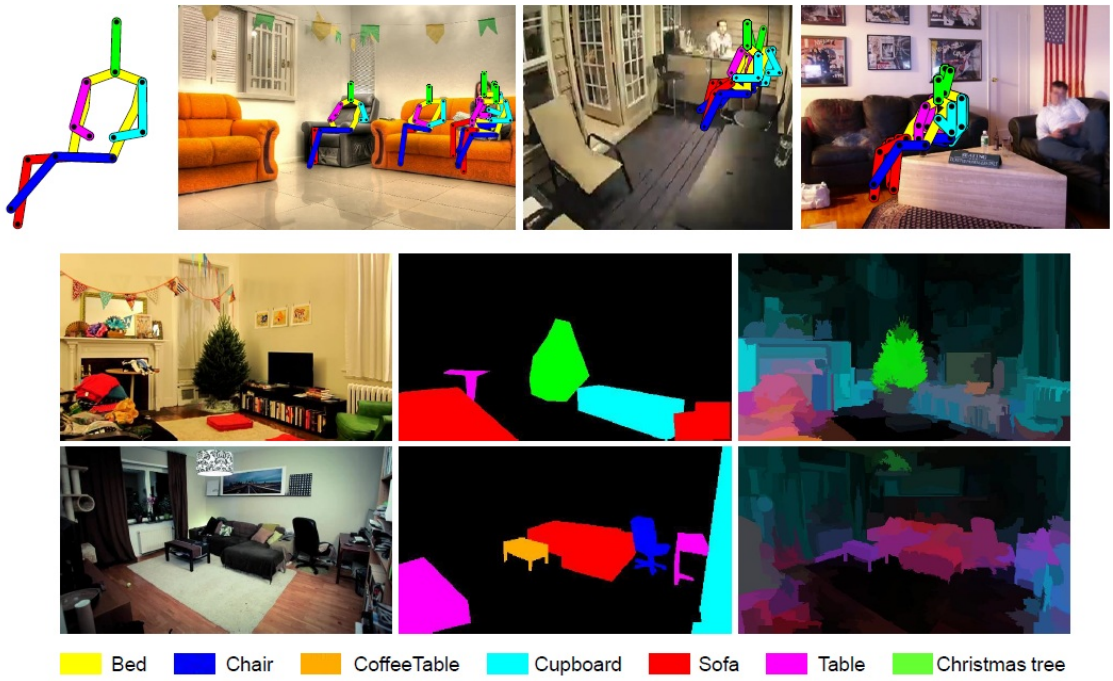

Figure 3. *Top: Example of particular pose detections in three indoor scenes. Bottom: object segmentation illustrated by original images, ground truth segmentation, and automatic segmentation by our method shown in the left, middle and right columns respectively.*

6.4.2. Analysis of Crowded Scenes in Video

Participants: Ivan Laptev, Josef Sivic, Mikel Rodriguez [MITRE].

In this work we first review the recent studies that have begun to address the various challenges associated with the analysis of crowded scenes. Next, we describe our two recent contributions to crowd analysis in video. First, we present a crowd analysis algorithm powered by prior probability distributions over behaviors that are learned on a large database of crowd videos gathered from the Internet. The proposed algorithm performs like state-of-the-art methods for tracking people having common crowd behaviors and outperforms the methods when the tracked individuals behave in an unusual way. Second, we address the problem of detecting and tracking a person in crowded video scenes. We formulate person detection as the optimization of a joint energy function combining crowd density estimation and the localization of individual people. The proposed methods are validated on a challenging video dataset of crowded scenes. Finally, the chapter concludes by describing ongoing and future research directions in crowd analysis.

This work is to appear in [17].

6.4.3. Actlets: A Novel Local Representation for Human Action Recognition in Video

Participants: Muhammad Muneeb Ullah, Ivan Laptev.

This work addresses the problem of human action recognition in realistic videos. We follow the recently successful local approaches and represent videos by means of local motion descriptors. To overcome the huge variability of human actions in motion and appearance, we propose a supervised approach to learn local motion descriptors – *actlets* – from a large pool of annotated video data. The main motivation behind our method is to construct action-characteristic representations of body joints undergoing specific motion patterns while learning invariance with respect to changes in camera views, lighting, human clothing, and other factors. We avoid the prohibitive cost of manual supervision and show how to learn actlets automatically from synthetic videos of avatars driven by the motion-capture data. We evaluate our method and show its significant improvement as well as its complementarity to existing techniques on the challenging UCF-sports and YouTube-actions datasets.

This work has been published in [16].

6.4.4. Layered Segmentation of People in Stereoscopic Movies

Participants: Karteek Alahari, Guillaume Seguin, Josef Sivic, Ivan Laptev.

In this work we seek to obtain a layered pixel-wise segmentation of multiple people in a stereoscopic video. This involves challenges such as dealing with unconstrained stereoscopic video, non-stationary cameras, complex indoor and outdoor dynamic scenes. The contributions of our work are three-fold: First, we develop a layered segmentation model incorporating person detections and pose estimates, as well as colour, motion, and stereo disparity cues. The model also explicitly represents depth ordering and occlusions of people. Second, we introduce a stereoscopic dataset with frames extracted from feature length movies “StreetDance 3D" and “Pina". In addition to realistic stereo image data, it contains nearly 700 annotated poses, 1200 annotated detections, and 400 pixel-wise segmentations of people. Third, we evaluate the benefits of stereo signal for person detection, pose estimation and segmentation in the new dataset. We demonstrate results on challenging realistic indoor and outdoor scenes depicting multiple people with frequent occlusions. Example result is shown in Figure 4 .

This work has been submitted to CVPR 2013.

6.4.5. Highly-Efficient Video Features for Action Recognition and Counting

Participants: Vadim Kantorov, Ivan Laptev.

Figure 4. A sample frame extracted from the stereoscopic movie "StreetDance": From left to right – left image from the stereo pair, disparity map computed from the stereo pair, layered segmentation of the image into 7 people. The front to back ordering is shown as a colour map, where "blue" denotes front and "red" denotes back. The cost function associated with our model is initialized using person detections, and incorporates disparity, pose, colour and motion cues. Note that the result shows accurate segmentation boundaries and also a reliable layer ordering of people.

Local video features provide state-of-the-art performance for action recognition. While the accuracy of action recognition has been steadily improved over the recent years, the low speed of feature extraction remains to be a major bottleneck preventing current methods from addressing large-scale applications. In this work we demonstrate that local video features can be computed very efficiently by exploiting motion information readily-available from standard video compression schemes. We show experimentally that the use of sparse motion vectors provided by the video compression improves the speed of existing optical-flow based methods by two orders of magnitude while resulting in limited drops of recognition performance. Building on this representation, we next address the problem of event counting in video and present a method providing accurate counts of human actions and enabling to process 100 years of video on a modest computer cluster.

This work has been submitted to CVPR 2013.